SFB 649 Discussion Paper 2016-050

Network Quantile Autoregression

Xuening Zhu*
Weining Wang *²
Hangsheng Wang*
Wolfgang K. Härdle*²

* Peking University, People's Republic of China;
*² Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Network Quantile Autoregression

Xuening Zhu[2], Weining Wang[1,3], Hangsheng Wang[2], and Wolfgang Karl Härdle[1,4]

[1] *Humboldt-Universität zu Berlin, C.A.S.E.-Center of Applied Statistics and Economics, Unter den Linden 6, 10099 Berlin, Germany;*

[2] *Guanghua School of Mangement, Peking University, Beijing, China;*

[3] *Department of Economics, City University London;*

[4] *Research fellow in Sim Kee Boon Institute for Financial Economics, Singapore Management University, 90 Stamford Road, 6th Level, School of Economics, Singapore* 178903.

Abstract

It is a challenging task to understand the complex dependency structures in an ultra-high dimensional network, especially when one concentrates on the tail dependency. To tackle this problem, we consider a network quantile autoregression model (NQAR) to characterize the dynamic quantile behavior in a complex system. In particular, we relate responses to its connected nodes and node specific characteristics in a quantile autoregression process. A minimum contrast estimation approach for the NQAR model is introduced, and the asymptotic properties are studied. Finally, we demonstrate the usage of our model by investigating the financial contagions in the Chinese stock market accounting for shared ownership of companies.

KEY WORDS: Social Network; Quantile Regression; Autoregression; Systemic Risk; Financial Contagion; Shared Ownership.

*This research is supported in part by National Natural Science Foundation of China (NSFC, 11131002, 11271032) and Center for Statistical Science at Peking University. This research is partly supported by DFG (WA3868/1-1) " quantile method for complexed financail system". We acknowledge the support from German Research Foundation DFG on SFB649.

1. INTRODUCTION

Quantile regression is an increasingly popular tool for modern statistical analysis. Instead of studying the conditional mean function of the response variable, quantile regression is concerned with estimating the conditional quantile function. It has been applied to a wide range of applications, such as labor economics (Koenker and Hallock, 2001; Fitzenberger et al., 2013), financial risk management (Gaglianone et al., 2012; Härdle et al., 2016), and environmental statistics (Sankarasubramanian and Lall, 2003; Wang et al., 2013). Particularly, the linear quantile model has been studied by a seminal work by Koenker and Bassett (1978), and the asymptotic theory has been developed by Portnoy (1991, 1997). See Koenker (2005) for a comprehensive summary of the methods and applications.

Following the well development of quantile method in the existing literature, the quantile regression in time series is of particular interest. An early stream of research such as Hallin et al. (1999), Hasan and Koenker (1997) deal with linear quantile autoregression model, which focus on independent identically distributed (*iid*) innovations in a relative restrictive location shift model. As another approach, Engle and Manganelli (2004) propose a set of autoregressive forms (referred to as CaViaR model) for value-at-risk (VaR), which all require to estimate VaR at first step and plug in a prescribed time series model. The framework is easy to apply but is quite difficult to directly infer the underlying process. As an alternative, Koenker and Xiao (2006) develop a quantile autoregressive method to model the conditional quantile function, which allows to study the asymptotic properties of the underlying process and does not assume an *iid* underlying process. However, it can only be employed in the univariate case. There have been a few efforts to develop multivariate quantile time series model. White et al. (2010) propose a generalization of the CaViaR method to a vector autoregression model for VaR. Recently, Baruník and Kley (2015) consider the quantile cross-spectral analysis of quantile vector autoregression processes. However, to our best knowledge, the existing methods are not flexible to directly extend to high dimensional data, as the number of parameters becomes large. This makes the estimation not only suboptimal, but also with high computational complexity. Therefore, a tractable quantile autoregression model for analyzing high dimensional data needs to be developed.

In the meanwhile, the rapid development of modern computer science and technology has allowed us to approach large amount of data with network structure. On one hand, it poses challenges on analyzing the dynamic processes with high dimensions. On the other hand, this brings us a unique opportunity to develop network models with network information naturally embedded. To take advantage of this extremely valuable information (i.e., network structure), we established a network quantile autoregression model (NQAR), which allows us to both make inference on the underlying processes and handle high dimensional modeling issues.

In the existing literature, great efforts have been taken to incorporate the network information into the modeling framework. For instance, Sewell and Chen (2015) take advantage of the network information to study the dynamic social behavior of students in a Dutch class by a latent space model. The community detection and extraction methods are studied by Zhao et al. (2011), Amini et al. (2013), and Sewell et al. (2016) using the block network structure. Accordingly, the corresponding consistency properties are established (Bickel and Chen, 2009; Zhao et al., 2012). Recently, the autoregression models in large-scale social networks receive great attention, where the estimation and computation issues are intensively discussed (Zhang and Chen, 2013; Zhou et al., 2015; Huang and Wang, 2016; Zhu et al., 2016). For other related statistical methods and applications, see Carrington et al. (2005), Newman (2010), and Kolaczyk and Csárdi (2014) for a comprehensive summary.

However, to our best knowledge, none of these aforementioned works have taken account the network information into a quantile regression framework. In this work, we provide an innovative network quantile autoregression model to better estimate and predict conditional quantiles in complex network systems (e.g. a network of stocks in a stock market). The model is based on a univariate quantile autoregression model. It is assumed the conditional quantile function of the response variable (e.g., volatility of stock returns) is related to several terms. The included terms are nodal specific variables (e.g., firm specific variables), the lagged response of the same node (e.g., volatility of the same stock in the previous time point), and the lagged responses of the other connected nodes (e.g., volatility of the related stocks in the previous time points). In our application, the connection between nodes are decided by the network information from the data, i.e., if two stocks share the same shareholder, then they are defined as connected.

Our paper contributes to the literature in three aspects. Firstly, we provide a network quantile model that characterizes the dynamic quantile behavior for high dimensional processes. Secondly, we propose a model framework to incorporate valuable network information from the data. Thirdly, the asymptotic theories are derived for both the underlying processes and estimated parameters.

The rest of the paper is organized as follows. Section 2 introduces the network quantile autoregression model, where the stationary results are established. Section 3 proposes a novel impulse analysis framework for the network quantile autoregression model. The parameter estimation method is given in Section 4, where the asymptotic properties are given. Extensive numerical studies and a real data analysis for stocks in Chinese financial markets are conducted in Section 5. Lastly, a brief conclusion is discussed in

Section 6. All technical details are delegated to the appendix.

2. NETWORK QUANTILE AUTOREGRESSION

2.1. Model and Notations

Let U_{it} ($1 \le i \le N, 1 \le t \le T$) be a sequence of *iid* random variables, which follows the standard uniform distribution. Assume that a q-dimensional random nodal covariate vector $Z_i \in \mathbb{R}^q$ is collected for each node. To record the network relationship, we define $A = (a_{ij}) \in \mathbb{R}^{N\times N}$ as the adjacency matrix, where $a_{ij} = 1$ if there is an edge from i to j, otherwise $a_{ij} = 0$. Following the convention, the nodes are assumed to be not self-related (i.e., $a_{ii} = 0$). Motivated by the univariate autoregression quantile model (Koenker and Xiao, 2006), we consider the network quantile autoregression (NQAR) model as

$$Y_{it} = \beta_0(U_{it}) + \sum_{l=1}^{q} Z_{il}\gamma_l(U_{it}) + \beta_1(U_{it})n_i^{-1}\sum_{j=1}^{N} a_{ij}Y_{i(t-1)} + \beta_2(U_{it})Y_{i(t-1)} \stackrel{\text{def}}{=} g_\theta(U_{it}), \quad (2.1)$$

where β_js ($0 \le j \le 2$) and γ_ls ($1 \le l \le q$) are unknown coefficient functions from $[0, 1]$ to $\mathbb{R}^1$, and $n_i = \sum_{j\neq i} a_{ij}$ is the out-degree for the ith node.

Denote $Q_Y(\tau|X)$ as the conditional quantile function of Y given X. By assuming the right side of (2.1) is monotonically increasing in U_{it}, we can write the conditional quantile function of Y_{it} given $(Z_i, \mathbb{Y}_{t-1})$ as:

$$Q_{Y_{it}}(\tau|Z_i, \mathbb{Y}_{t-1}) = \beta_0(\tau) + \sum_{l=1}^{q} Z_{il}\gamma_l(\tau) + \beta_1(\tau)n_i^{-1}\sum_{j=1}^{N} a_{ij}Y_{j(t-1)} + \beta_2(\tau)Y_{i(t-1)}.$$

It is worth mentioning that the varying coefficients are functions of τ. Therefore, not only the location of the conditional density of Y_{it} is determined by τ, but also the shape of the conditional density is allowed to be τ-dependent. Specifically, $\beta_0(\tau) + \sum_{l=1}^{q} Z_{il}\gamma_l(\tau)$ reflects the nodal impact with respect to node i, where $\beta_0(\tau)$ is referred to as the *baseline function*. The covariates Z_{il} refer to node-specific variables, which are invariant over t. For example, it can be stock related features (e.g., size, leverage ratio). It is assumed the nodal covariates Z_is are independent of the U_{it}s. Next, the second term $n_i^{-1}\sum_{j=1}^{N} a_{ij}Y_{j(t-1)}$ characterizes the network impact from the connected nodes (e.g., stocks with same shareholders). The corresponding coefficient function $\beta_1(\tau)$ is then referred to as the *network function*. Lastly, $Y_{i(t-1)}$ captures the impact from the response of the same node in the previous time point. Accordingly, the coefficient function $\beta_2(\tau)$ is then referred to as the *momentum function*. To obtain more insights of the NQAR model, we discuss the model under the following three scenarios.

Scenario 1. (Tail Behavior) The NQAR model can capture the asymmetric effect between the responses at different quantile levels. In particular, it is of great interest to understand the tail dependency of the responses. For instance, to model the conditional VaR of the stock return, one could define Y_{it} to be the return for the ith stock and fix $\tau = 0.05$ for analysis. In such a situation, an asymmetric pattern gives indication on whether the financial institutions tend to have closer connections in the lower tail (e.g. in the financial crisis) than at the other levels (e.g., median level and upper tail).

Scenario 2. (Robust Estimation) In many occasions, the estimation can be seriously distorted by outliers in the dataset (Abello et al., 2013; Li et al., 2015). To obtain reliable results, robust estimation can be conducted. Compared to vector autoregression for the mean case (Lütkepohl, 2005), the NQAR is insensitive extreme values. In particular, robust median estimation can be obtained by setting $\tau = 0.5$.

Scenario 3. (Constant Coefficient Function) Consider the case that the coefficient functions exhibit constant forms, i.e. $\beta_j(\tau) = \beta_j$ $(j = 1, 2)$ and $\gamma_l(\tau) = \gamma_l$ $(l = 1, \cdots, q)$. In such a situation, the conditional distribution of Y_{it} is τ-invariant. We further let $\beta_0(u) = \sigma\Phi^{-1}(u)$, where $\Phi(\cdot)$ is the distribution function of the standard normal distribution. Then the NQAR model degenerates to the network autoregression (NAR) model in the mean case, see Zhu et al. (2016).

For convenience, define $\mathbb{Y}_t = (Y_{1t}, \cdots, Y_{Nt})^\top \in \mathbb{R}^N$, $\mathbb{Z} = (Z_1, \cdots, Z_N)^\top \in \mathbb{R}^{N\times q}$. Let $\mathbf{B}_{0t} = \big(\beta_0(U_{it}) + \sum_l Z_{il}\gamma_l(U_{it}), 1 \le i \le N\big)^\top \in \mathbb{R}^N$, $\mathbf{B}_{1t} = \mathrm{diag}\{\beta_1(U_{it}), 1 \le i \le N\} \in \mathbb{R}^{N\times N}$, $\mathbf{B}_{2t} = \mathrm{diag}\{\beta_2(U_{it}), 1 \le i \le N\}^\top \in \mathbb{R}^{N\times N}$. One easily verifies that $\Gamma = \mathsf{E}(\mathbf{B}_{0t}) = c_0\mathbf{1}_N \in \mathbb{R}^N$, where $c_0 = b_0 + c_Z$, $b_0 = \int_0^1 \beta_0(u)du$ and $c_Z = \mathsf{E}(Z_1)^\top r$ with $r = \big(\int_0^1 \gamma_l(u)du, 1 \le l \le q\big)^\top \in \mathbb{R}^q$. Without loss of generality, we set $\mathsf{E}(Z_1) = \mathbf{0}$. Then the NQAR model (2.1) can be re-written in vector form as

$$\mathbb{Y}_t = \Gamma + G_t\mathbb{Y}_{t-1} + V_t, \tag{2.2}$$

where $G_t = \mathbf{B}_{1t}W + \mathbf{B}_{2t} \in \mathbb{R}^{N\times N}$, $W = (w_{ij}) = (n_i^{-1}a_{ij}) \in \mathbb{R}^{N\times N}$ is the row-normalized adjacency matrix, and $V_t = \mathbf{B}_{0t} - \Gamma \in \mathbb{R}^N$ is *iid* with mean $\mathbf{0}$ and covariance $\Sigma_V = \sigma_V^2 I_N \in \mathbb{R}^{N\times N}$ with $\sigma_V^2 = \sigma_{b_0}^2 + \mathsf{E}\{\gamma^\top(U_{1t})\Sigma_Z\gamma(U_{1t})\}$, $\sigma_{b_0}^2 = \int_0^1 \beta_0^2(u)du - \{\int_0^1 \beta_0(u)du\}^2$, and $\Sigma_Z = \mathsf{Cov}(Z_1) \in \mathbb{R}^{q\times q}$.

2.2. Covariance Stationarity

Given the NQAR model (2.2), it is then of great interest to check stationarity of $\mathbb{Y}_t$. A process $\{\mathbb{Y}_t\}_{-\infty}^{+\infty}$ is covariance stationary if (a) $\mathsf{E}(\mathbb{Y}_t) = \mu_Y$ for a constant vector $\mu_Y \in \mathbb{R}^N$; (b) $\mathsf{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-h}) = \mathsf{E}\{(\mathbb{Y}_t - \mu_y)(\mathbb{Y}_{t-h} - \mu_y)^\top\} = \Sigma(h)$ with $\Sigma(h) \in \mathbb{R}^{N\times N}$ only related to h. For convenience, let $b_1 = \mathsf{E}\{\beta_1(U_{it})\}$, $b_2 = \mathsf{E}\{\beta_2(U_{it})\}$, $\widetilde{b}_1 = \{\mathsf{E}\,\beta_1^2(U_{it})\}^{1/2}$, $\widetilde{b}_2 = \{\mathsf{E}\,\beta_2^2(U_{it})\}^{1/2}$, $G = \mathsf{E}(G_t)$, and $G^* = \mathsf{E}(G_t \otimes G_t)$. Then we have the following theorem.

THEOREM 2.1. *Assume* $\widetilde{b}_1 + \widetilde{b}_2 < 1$ *and* $\mathsf{E}\,|V_{it}| < C$ *for some positive constant* C.

(a) There exists a unique covariance stationary solution of the NQAR model (2.2) with finite first order moment as

$$\mathbb{Y}_t = \sum_{l=0}^{\infty} \Pi_l \Gamma + \sum_{l=0}^{\infty} \Pi_l V_{t-l}, \tag{2.3}$$

where $\Pi_l = \prod_{k=1}^{l} G_{t-k+1}$ *for* $l \geq 1$ *and* $\Pi_0 = I_N$.

(b) Denote $\Sigma_Y = \Sigma(0)$. *The stationary mean is* $\mu_Y = c_1^{-1} c_0 \mathbf{1}_N$ *and*

$$vec(\Sigma_Y) = (M_1 - c_1^{-2} c_0^2)\mathbf{1}_{N^2} + c_1^{-1} c_0 (I - G^*)^{-1} vec(\Sigma_{bv}) + (I - G^*)^{-1} vec(\Sigma_V), \tag{2.4}$$

where $c_1 = (1 - b_1 - b_2)^{-1}$, $M_1 = c_1^{-1} c_0^2 (1 + b_1 + b_2)(I - G^*)^{-1}$, $\Sigma_{bv} = \sigma_{bv} I_N$, *and* $\sigma_{bv} = \mathsf{E}[\{\beta_1(U_{it}) + \beta_2(U_{it})\} V_{it}]$. *Moreover, we have* $\Sigma(h) = G^h \Sigma_Y$ *for any integer* $h > 0$, *and* $\Sigma(h) = \Sigma_Y (G^\top)^{-h}$ *for* $h < 0$.

The proof of Theorem 2.1 is given in Appendix A.1. By Theorem 2.1, unique covariance stationary solution (2.3) of the NQAR model is given. To obtain more insights of Theorem 2.1, we would like to add two remarks.

Remark 1. It is straightforward to verify $\widetilde{b}_1 = (b_1^2 + \sigma_{b_1}^2)^{1/2}$, where $\sigma_{b_1}^2 = \mathsf{Var}\{\beta_1(U_{it})\}$. Similarly one can define $\sigma_{b_2}^2 = \mathsf{Var}\{\beta_2(U_{it})\}$. Therefore the stationarity assumption in Theorem 2.1 essentially sets constraint on the expectation and variance of the network and momentum functions (i.e., $\beta_1(\cdot)$ and $\beta_2(\cdot)$). It is noteworthy that the assumption does not require $|\beta_1(\tau)| + |\beta_2(\tau)|$ to be strictly less than one for every τ. Instead, it imposes an upper bound on average levels, which allows for some "explosive" cases in a specific quantile (i.e., $|\beta_1(\tau)| + |\beta_2(\tau)| > 1$ for some τ). Particularly, if the network and momentum functions take constant forms, i.e., $\beta_1(\tau) = b_1$ and $\beta_2(\tau) = b_2$ for some constants b_1 and b_2, then the stationary assumption reduces to $|b_1| + |b_2| < 1$, which corroborates to the stationary condition in the mean case (Zhu et al., 2016).

Remark 2. Let us look at the stationary mean μ_y and covariance Σ_Y. First, note that $\mu_Y = c_1^{-1}c_0\mathbf{1}_N$, is the same for all the nodes, and unrelated to the network structure. However, the situation for the covariance Σ_Y is more complicated. The network effect though tends to be very small (Chen et al., 2013; Zhu et al., 2016). This motivates us to assume that $\widetilde{b}_1 = o(1)$, and then approximate the Σ_Y by the leading terms. For convenience, define $\widetilde{b}_{12} = \mathsf{E}\{\beta_1(U_{it})\beta_2(U_{it})\}$, $\widetilde{b}_{01} = \mathsf{E}\{\beta_1(U_{it})V_{it}\}$, and $\widetilde{b}_{02} = \mathsf{E}\{\beta_2(U_{it})V_{it}\}$. Then we have,

$$\mathsf{Var}(Y_{it}) \approx c_{b1}c_0^2 + (1-\widetilde{b}_2^2)^{-1}\{2(1-b_2)^{-2}\{(1-b_2)\sigma_{bv} + b_1\widetilde{b}_{02}\}c_0 + \sigma_V^2\}, \tag{2.5}$$

$$\mathsf{Cov}(Y_{i_1t}, Y_{i_2t}) \approx c_{b2}c_0^2 + (1-b_2^2)^{-2}\{2(1-b_2)^{-1}\widetilde{b}_{02}c_0 + \sigma_V^2\}b_1b_2(w_{i_1i_2} + w_{i_2i_1}), \tag{2.6}$$

where $c_{b1} = [(1-\widetilde{b}_2^2)^{-1}\{1-b_2^2+2b_1+2(1-\widetilde{b}_2^2)^{-1}(1-b_2^2)\widetilde{b}_{12}\}-(1-b_2)^{-1}(1-b_2+2b_1)](1-b_2)^{-2}$ and $c_{b2} = (1-b_2)^{-2}(1-b_2^2)^{-2}(1-b_2^2+2b_1+2b_1b_2)-(1-b_2)^{-3}(1-b_2+2b_1)$. The verifications of (2.5) and (2.6) are given in Appendix A.2. It is worth noting that the variance of Y_{it} is mainly determined by the momentum function $\beta_2(\cdot)$ and the baseline function $\beta_0(\cdot)$. Particularly, larger $\widetilde{b}_2$ will result in higher variance of Y_{it}. Next, the covariance between Y_{i_1t} and Y_{i_2t} is not only related to $\beta_2(\cdot)$, but is also related to the network function $\beta_1(\cdot)$. Nodes have a higher correlation with each other if b_1 is larger. Note that $w_{i_1i_2}+w_{i_2i_1} = n_{i_1}^{-1}a_{i_1i_2}+n_{i_2}^{-1}a_{i_2i_1}$. Therefore, the correlation between node i_1 and i_2 is higher if (a) they connect to each other in the network (i.e., $a_{i_1i_2} = a_{i_2i_1} = 1$) and (b) they both have small out-degrees (i.e., small n_{i_1} and n_{i_2}).

2.3. Asymptotic Stationary Distribution

Given the established covariance stationarity, it is then natural to derive the asymptotic stationary distribution. However, it is not straightforward to derive the asymptotic stationary distribution of $\mathbb{Y}_t$. Instead, we focus on the long run average of $\mathbb{Y}_t$. Accordingly, let $\overline{\mathbb{Y}}_T = T^{-1} \sum_{t=1}^{T} \mathbb{Y}_t$ be the average of the responses during T periods, we then have the following theorem.

THEOREM 2.2. *Assume $c_\beta < 1$ and $E(|V_{it}|^4) < M$, where $c_\beta = \|\beta_1\|_4 + \|\beta_2\|_4$ with $\|\beta_j\|_4 = \mathsf{E}\{\beta_j(U_{it})^4\}^{1/4}$ $(j = 1, 2)$, and M is finite positive constant. Then the covariance stationary solution in (2.3) follows the central limit theorem as*

$$\sqrt{T}(\overline{\mathbb{Y}}_T - \mu_Y) \xrightarrow{\mathcal{L}} \mathbb{N}(0, \Sigma_Y) \tag{2.7}$$

as $T \to \infty$.

The proof of Theorem 2.2 is given in Step 3 Appendix A.1. By (2.7), it can be seen that the asymptotic covariance of $\sqrt{T}(\overline{\mathbb{Y}}_T - \mu_Y)$ is equivalent to $\mathsf{Cov}(\mathbb{Y}_t) = \Sigma_Y$.

3. IMPULSE ANALYSIS

3.1. Measurements of Impulse Effect

In this section, we aim to investigate how a node in the network will react to an exogenous shock imposed on the other nodes, which is referred to as an impulse analysis. Particularly, consider a stimulus $\Delta = (\delta_1, \cdots, \delta_N)^\top \in \mathbb{R}^N$ imposed on V_t, and shock it to $V_t + \Delta$. Then, the response for the ith node at time point t (i.e., Y_{it}) will grow to $Y_{it} + \delta_i$. Following the NQAR model (2.2), the response at time point $(t + l)$, $l \geq 1$ (i.e., $\mathbb{Y}_{t+l}$) is increased by

$$\mathrm{IE}_{t,t+l} = \prod_{k=0}^{l-1} G_{t+l-k} \Delta, \tag{3.1}$$

where $\mathrm{IE}_{t,t+l}$ is referred to as the *impulse effect* from time t to $t + l$. For instance, if $\Delta = (1, 0, \cdots, 0)^\top$, then the $\mathrm{IE}_{t,t+l}$ is the first column of $\prod_{k=0}^{l-1} G_{t+l-k}$. It can be noted that the impulse effect is only related to the network and momentum functions, i.e., $\beta_1(\cdot)$ and $\beta_2(\cdot)$.

However, by this definition of (3.1), $\mathrm{IE}_{t,t+l}$ is a random vector so that it cannot be directly calculated. To fix the problem, the following three measurements of impulse effects are considered.

1. Average Impulse Effect. The average impulse effect (AIE) is defined as the expectation of $\mathrm{IE}_{t,t+l}$ as $\mathsf{E}(\mathrm{IE}_{t,t+l}) = G^l\Delta = (b_1W + b_2I_N)^l\Delta$. Specifically, the AIE is only related to the average network and momentum effect b_1 and b_2. Furthermore, it is noteworthy that the AIE is no longer related to t but only depends on the time difference l. It can be further derived$|\mathbf{1}^\top \mathsf{E}(\mathrm{IE}_{t,t+l})| \leq N(\widetilde{b}_1 + \widetilde{b}_2)^l C_\Delta$, where $C_\Delta = \max_i |\Delta_i|$. Therefore, it can be concluded that the total network AIE will decrease to 0 as $l \to \infty$, if the stationary condition in Theorem 2.1 is satisfied.

2. Interval Impulse Effect. Although the AIE can characterize the impulse effect on an average level, it is hard to capture the asymmetric effect for different quantiles. To this end, we define the interval impulse effect (IIE) from t to $t+l$ within the interval $[\tau_1, \tau_2]$, $(0 \leq \tau_1 < \tau_2 \leq 1)$ as

$$\begin{aligned}
\mathrm{IIE}_{l,\tau_1\tau_2} &= \mathsf{E}\Big\{\prod_{k=0}^{l-1} G_{t+l-k}\Delta \Big| U_{im} \in [\tau_1, \tau_2], 1 \leq i \leq N, t+1 \leq m \leq t+l\Big\} \\
&= (c_{\beta_1,\tau_1\tau_2}W + c_{\beta_2,\tau_1\tau_2}I_N)^l\Delta,
\end{aligned}$$

where $c_{\beta_1,\tau_1\tau_2} = \int_{\tau_1}^{\tau_2} \beta_1(u)du$ and $c_{\beta_2,\tau_1\tau_2} = \int_{\tau_1}^{\tau_2} \beta_2(u)du$. As one can see, the amount of IIE is determined by the integration of $\beta_1(u)$ and $\beta_2(u)$ within the interval $[\tau_1, \tau_2]$. Note that the interval $[\tau_1, \tau_2]$ can be any interested regions. For example, to measure the asymmetric effects of the upper tail, median level, and the lower tail, one can split $(0, 1)$ equally into three intervals (i.e., $[0, 1/3)$, $[1/3, 2/3)$, $[2/3, 1]$) and compare the

IIEs for different intervals respectively.

3. Impulse Effect Intensity. By the IIE, the asymmetric effects can be readily quantified. However, due to the unknown function form of β_1 and β_2, the integration can still be hard to compute. On the other hand, note the fact that the IIE can be defined in arbitrary intervals in $[0, 1]$. Motivated by this, we consider a sufficiently small interval $[\tau, \tau + \delta]$, and define the impulse effect intensity (IEI) at τ as

$$\begin{aligned}\mathrm{IEI}_{l,\tau} &= \lim_{\delta\to 0} \delta^{-l} E\Big\{ \prod_{k=0}^{l-1} G_{t+l-k}\Delta \Big| U_{im} \in [\tau, \tau+\delta], 1 \le i \le N, t+1 \le m \le t+l \Big\} \\ &= \big\{\beta_1(\tau)W + \beta_2(\tau)I_N\big\}^l \Delta,\end{aligned}$$

where $\beta_1(u)$ and $\beta_2(u)$ are assumed to be continuous at τ. By this definition, $\mathrm{IEI}_{l,\tau}$ could reflect impulse impact at the τth quantile. Moreover, the quantity $\mathrm{IEI}_{l,\tau}$ is easy to compute once the estimates of $\beta_1(\tau)$ and $\beta_2(\tau)$ are obtained.

Given the three types of impulse effect measurement, the cross-sectional impulse analysis can be conducted. Assume that one unit stimulus is imposed on the ith node, a cross-sectional impulse analysis is about analyzing its impact on the other nodes. This kind of analysis is interesting in a network of banks. It delivers an importance message on the systemic risk spillover of an institution. Take the IEI as an example and assume $\Delta = (\delta_i)^\top$ with only $\delta_i = 1$ and $\delta_{i'} = 0$ (for all $i' \ne i$). The IEI from node i to j can be defined by the jth element of $\mathrm{IEI}_{l,\tau}$, which is then denoted as $\mathrm{IEI}_{l,\tau}^{(i,j)}$. Equivalently, $\mathrm{IEI}_{l,\tau}^{(i,j)}$ is equal to the (j,i)th element of the matrix $\big\{\beta_1(\tau)W + \beta_2(\tau)I_N\big\}^l$. Then, if $\mathrm{IEI}_{l,\tau}^{(i,j)}$ is sufficiently large and decays slowly as $l \to \infty$, then the jth risk factor can be seriously affected by the fluctuations of the ith risk factor. Lastly, it is worth mentioning that the cross-sectional impulse analysis can also be conducted by using the other two impulse effect measurements, which can be defined in the similar way with details ignored here.

3.2. Influential Node Analysis

It is worth noting that the impulse effect can be interpreted as the instantaneous impact from t to $t+l$. Moreover, we define the total network average impulse effect (TNAIE) as the summation of AIE over all nodes and all horizons l as $\mathrm{TNAIE}(\Delta) = \sum_{l=0}^{\infty} \mathbf{1}^\top \mathsf{E}(\mathrm{IE}_{t,t+l}) = \sum_{l=0}^{\infty} \mathbf{1}^\top G^l \Delta = \mathbf{1}^\top (I_N - G)^{-1}\Delta$. It should be noted that the definition of TNAIE can also be extended using the other two impulse effect measures (IIE and IEI). For the sake of their similarities, we skip the details and use the AIE-defined TNAIE in this section.

Write TNAIE(Δ) as $\mathrm{TNAIE}(\Delta) = \sum_{i=1}^{N} \nu_i \delta_i$, where ν_i is the ith element of the vector $\nu = (I - G^\top)^{-1}\mathbf{1}$. Consider that one unit stimulus is imposed on the node i, that $\delta_i = 1$ and $\delta_j = 0$ for all $j \neq i$. Then we have $\mathrm{TNAIE}(\Delta) = \delta_i$ as a result. This reflects the amount that the whole network will react to the unit perturbation of the node i. For convenience, we refer to ν_i as the *influential power* of node i.

As one can see, the definition of ν_i is straightforward but it can be hard to compute. This is because the calculation of ν involves the inverse of a high dimensional matrix $(I - G^\top)$. Following the idea of Remark 2 of Theorem 2.1, we approximate the $\nu_i$values by first order Taylor expansion: $\nu_i \approx 1/(1-b_2)+(1-b_2)^{-2} b_1 \sum_j n_j^{-1} a_{ji}$. Assume $b_1 > 0$, then the influential power of node i is mainly determined by the quantity $\sum_j n_j^{-1} a_{ji}$, which is referred to as the *weighted degree* of the node i. Consequently, nodes with larger weighted degrees tend to be followed by a large amount of nodes (i.e., $\sum_j a_{ji}$). Moreover, at the same time, the followers should have less out-degrees (i.e., small n_js).

4. PARAMETER ESTIMATION

In this section, we provide an estimation method to the NQAR model (2.1). The asymptotic properties are also established. Write $\theta(\tau)$ to be a $(q+3)$-dimensional parameter vector as $\theta(\tau) = [\beta_0(\tau), \gamma^\top(\tau), \beta_1(\tau), \beta_2(\tau)]^\top \in \mathbb{R}^{p+3}$. Define $g_{\theta,it}(\tau) = X_{it}^\top \theta(\tau)$, and $V_{it\tau} = Y_{it} - g_{\theta,i(t-1)}(\tau)$ where $X_{it} = (1, Z_i^\top, n_i^{-1}\sum_{j=1}^N a_{ij}Y_{jt}, Y_{it})^\top \in \mathbb{R}^{q+3}$. Then the parameter vector $\theta(\tau)$ can be estimated by

$$\widehat{\theta}(\tau) = \arg\min_\theta \sum_{i=1}^N \sum_{t=1}^T \rho_\tau\{y_{it} - g_{\theta,i(t-1)}(\tau)\}, \tag{4.1}$$

where $\rho_\tau(u) = u\{\tau - \mathbf{1}(u<0)\}$ is the check function for quantile regression.

Let the conditional density function of Y_{it} given $\mathcal{F}_{t-1}$ be $f_{i(t-1)}(\cdot)$. To facilitate the study of the asymptotic properties of our estimation, we define $\widehat{\Omega}_0 = (NT)^{-1}\sum_{i=1}^N \sum_{t=1}^T X_{it}X_{it}^\top$ and $\widehat{\Omega}_1 = (NT)^{-1}\sum_{i=1}^N \sum_{t=1}^T f_{it}\big(X_{it}^\top\theta(\tau)\big)X_{it}X_{it}^\top$ for any given $\tau \in (0,1)$. Specifically, $f_{it}\big(X_{it}^\top\theta(\tau)\big)$ can be estimated by $\hat{f}_{it}\big(X_{it}^\top\hat{\theta}(\tau)\big) = \{X_{it}^\top(\hat{\theta}(\tau_l) - \hat{\theta}(\tau_{l-1}))\}^{-1}(\tau_l - \tau_{l-1})$, where $\tau \in [\tau_{l-1}, \tau_l]$ and $\{\tau_l\}$ is an appropriately chosen sequence. Next, to prove the asymptotic properties of the estimated parameters, the following assumptions are required.

(C1) (Moment Assumption) Assume $c_\beta < 1$, where c_β is defined in Theorem 2.2. Further, assume that Z_is are independent and identically distributed random vectors, with mean 0 and covariance $\Sigma_z \in \mathbb{R}^{p\times p}$. Furthermore, its fourth order moment is finite. The same assumption is also needed for V_{it} across both $1 \le i \le N$ and $0 \le t \le T$. Moreover, we need $\{Z_i\}$ and $\{U_{it}\}$ to be mutually independent.

(C2) (Network Structure)

(C2.1) (Connectivity) Let the set of all the nodes $\{1, \cdots, N\}$ be the state space of a Markov chain, with the transition probability given by W. It is assumed the Markov chain is irreducible and aperiodic. In addition, define $\pi =$

$(\pi_i)^\top \in \mathbb{R}^N$ to be the stationary distribution vector of the Markov chain (i.e., $\pi_i \geq 0$, $\sum_i \pi = 1$, and $W^\top \pi = \pi$). It is assumed that $\sum_{i=1}^N \pi_i^2 \to 0$ as $N \to \infty$.

(C2.2) (Sparsity) Assume $|\lambda_1(W^*)| = \mathcal{O}(\log N)$, where W^* is defined to be a symmetric matrix as $W^* = W + W^\top$.

(C3) (Convergence) Assume $\widehat{\Omega}_1 \to_p \Omega_1$ as $N \to \infty$, where $\Omega_1 = (\Omega_{1,ij}) \in \mathbb{R}^{N\times N}$ is a positive definite matrix. In addition, assume the following limits exist. They are, respectively, $\kappa_1 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(\Sigma_Y)$, $\kappa_2 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(W\Sigma_Y)$, $\kappa_3 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(W\Sigma_Y W^\top)$, and $\kappa_4 = \lim_{N\to\infty} N^{-1}\mathrm{tr}\{(I-G)^{-1}\}$, $\kappa_5 = \lim_{N\to\infty} N^{-1}\mathrm{tr}\{W(I-G)^{-1}\}$. Here κ_j $(1 \leq j \leq 5)$ are fixed constants.

(C4) (Density) There exists positive constants $0 < c_1 < c_2 < \infty$ such that $c_1 \leq f_{it}(x) \leq c_2$ for all $1 \leq i \leq N, 1 \leq t \leq T$ with $x \in \mathbb{R}$.

(C5) (Monotonicity) It is assumed $\theta(\tau)^\top X_{it}$ $(1 \leq i \leq N, 1 \leq t \leq T)$ is monotone increasing functions with respect to τ.

We now comment on the conditions. Condition (C1) are standard conditions on the noise term V_{it}s, nodal covariates Z_is and $\beta(U_{it})$s for the parameter consistency results. This condition can be relaxed to allow for the weak dependence or mixing case over time. Condition (C2) is set for the network structure. Specifically, condition (C2.1) ensures a structure on the network connectivity, as it will result to all nodes in the network connecting each other within a finite number of steps, which corresponds to the empirical phenomenon named as "six degrees of separation". (C2.2) assures that the network structure is sufficiently sparse,i.e. the divergence rate of $\lambda_1(W^*)$ is slower than $\log(N)$. Condition (C3) is set on the design matrices. These are conditions needed to apply law of large number to the design matrices, which lead to a proper limit to the asymptotic covariance matrix. Importantly, it restricts the dependency between nodal covariates such that the convergence is ensured. Finally, condition (C4) requires the density function of the response distribution to be bounded from up and below.

Remark 3. The monotonicity assumption is imposed by condition (C5) to ensure the validness of the quantile regression. As it is mentioned in Koenker and Xiao (2006), the monotonicity of $X_{it}^\top \theta(\tau)$ is likely to be violated in some regions of the covariate space. We therefore refer to Koenker and Xiao (2006) and Fan and Fan (2010) as a discussion and solutions on the relevant issue.

The following theorem provides the consistency of the parameter estimation.

THEOREM 4.1. *Under condition (C1)–(C5), the following representation holds uniformly over $\tau \in B$ (i.e., B is a compact set in $(0,1)$),*

$$\widehat{\theta}(\tau) - \theta(\tau) = (NT)^{-1}\Sigma_\theta(\tau)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\psi_\tau(V_{it\tau}) + r_{NT}(\tau), \tag{4.2}$$

where $\Sigma_\theta(\tau) = \Omega_1^{-1}\Omega_0\Omega_1^{-1}$, $V_{it\tau} = Y_{it} - g_{\theta,i(t-1)}(\tau)$, and the remainder term satisfies $sup_{\tau\in B}|r_{NT}(\tau)| = o_p\big((NT)^{-1/2}\big)$. This leads to the consistency result that $\widehat{\theta}(\tau) \xrightarrow{p} \theta(\tau)$ as $\min\{N,T\} \to \infty$ uniformly for $\tau \in B$.

The proof of Theorem 4.1 is given in Appendix B. With the consistency of the parameter, we then present the asymptotic distribution of the estimated parameter.

THEOREM 4.2. *Under condition (C1)–(C5), we have*

$$\sqrt{NT}\Sigma_\theta^{-1/2}(\tau)\big\{\widehat{\theta}(\tau) - \theta(\tau)\big\} \xrightarrow{\mathcal{L}} B_{q+3}(\tau)$$

as $\min\{N,T\} \to \infty$, where $B_{q+3}(\tau)$ is a $(q+3)$-dimensional Brownian bridge, $\Sigma_\theta =$

$\Omega_1^{-1}\Omega_0\Omega_1^{-1}$ *with*

$$\Omega_0 = \begin{pmatrix} 1 & \mathbf{0}^\top & c_b & c_b \\ \mathbf{0} & \Sigma_z & \kappa_5\Sigma_z\overline{\gamma} & \kappa_4\Sigma_z\overline{\gamma} \\ c_b & \kappa_5\overline{\gamma}^\top\Sigma_z & \kappa_3 + c_b^2 & \kappa_2 + c_b^2 \\ c_b & \kappa_4\overline{\gamma}^\top\Sigma_z & \kappa_2 + c_b^2 & \kappa_1 + c_b^2 \end{pmatrix}, \tag{4.3}$$

$c_b = c_1^{-1}c_0$, *and* Ω_1 *is defined in condition (C3).*

The proof of Theorem 4.2 is given in Appendix B. To better understand the convergence result given in Theorem 4.2, we consider the case that for any fixed τ, that $B_{q+3}(\tau)$ reduces to $\mathbb{N}(0, \tau(1-\tau)I_{q+3})$. Specifically, we have the following Corollary.

COROLLARY 4.1. *Under condition (C1)–(C5), for any fixed* $\tau \in B$ *we have the result* $\sqrt{NT}\big\{\widehat{\theta}(\tau) - \theta(\tau)\big\} \xrightarrow{\mathcal{L}} \mathbb{N}(0, \tau(1-\tau)\Sigma_\theta(\tau))$ *as* $\min\{N,T\} \to \infty$, *where* $B \subset (0,1)$ *is a compact set.*

Corollary 4.1 is a direct implication from Theorem 4.2. By Corollary 4.1, the asymptotic normality can be obtained at aribitrary fixed τ. This enables us to conduct pointwise (for a fixed τ) inference on the estimated parameters. Specifically, the numerical details are given in the next section.

5. NUMERICAL STUDIES

5.1. Simulation Models

We consider three simulation settings in this subsection to illustrate the finite sample performance of the proposed NQAR model. The main difference lies in the generating mechanism of the network structure (i.e, A). The baseline, network, and the momentum function are set to be $\beta_0(u) = u$, $\beta_1(u) = 0.1\Phi(u)$, $\beta_2(u) = 0.4\{1 + \exp(u)\}^{-1}\exp(u)$. In addition, we fix the dimension of nodal covariates (i.e., Z_i) to be 5. The corresponding nodal functions are then set to be $\gamma_1(u) = 0.5\Phi(u)$, $\gamma_2(u) = 0.3\mathbb{G}(u, 1, 2)$, $\gamma_3(u) = 0.2\mathbb{G}(u, 2, 2)$, $\gamma_4(u) = 0.25\mathbb{G}(u, 3, 2)$, and $\gamma_5(u) = 0.2\mathbb{G}(u, 2, 1)$, where $\mathbb{G}(\cdot, a, b)$ is the Gamma distribution function with shape parameter a and scale parameter b.

To generate observations from the NQAR mechanism (2.1), the following procedures are performed. First, we generate u_{it}s $(1 \le i \le N, 1 \le t \le T)$ independently from a standard normal distribution $N(0, 1)$ and t-distribution with 5 degrees of freedom. Then, the random coefficients are obtained by substituting u_{it} into $\beta_j(\cdot)$ and $\gamma_l(\cdot)$ functions for $0 \le j \le 2$ and $1 \le l \le 5$. Next, the nodal covariates $Z_i = (Z_{i1}, \cdots, Z_{i5})^\top \in \mathbb{R}^5$ are sampled from a multivariate normal distribution $N(\mathbf{0}, \Sigma_z)$, where $\Sigma_z = (\sigma_{j_1 j_2})$ and $\sigma_{j_1 j_2} = 0.5^{|\gamma_1 - \gamma_2|}$. Lastly, we fix $\mathbb{Y}_0 = (1 - \hat{b}_1 - \hat{b}_2)^{-1}\hat{b}_0\mathbf{1}$, where $\hat{b}_j$ is the numerical mean of the $\beta_j(\cdot)$ function over a set of τs. Then $\mathbb{Y}_t$s are generated according to (2.1). We adopt three kinds of the adjacency matrix structures that are well-known in the literature. The details are given in the following.

Example 1. (Dyad Independence Model) Holland and Leinhardt (1981) introduce a Dyad Indenpendece Model with Dyad defined as $D_{ij} = (a_{ij}, a_{ji})$ for $1 \le i < j \le N$. It is assumed the different D_{ij}s are independent. Specifically, we set the probability of mutually connect dyads to be $P(D_{ij} = (1, 1)) = 20N^{-1}$ to ensure the network sparsity. Besides, set $P(D_{ij} = (1, 0)) = P(D_{ij} = (0, 1)) = 0.5N^{-0.8}$, which implies that the expected degree for each node is $\mathcal{O}(N^{0.2})$. Accordingly, we have $P(D_{ij} = (0, 0)) = 1 - 20N^{-1} - N^{-0.8}$, which is close to 1 as $N \to \infty$.

Example 2. (Stochastic Block Model) We further consider the stochastic model, which is extensively studied in network analysis (Wang and Wong, 1987; Nowicki and

Snijders, 2001). In particular, it is important for community detection (Zhao et al., 2012). To generate the block network structure, we follow Nowicki and Snijders (2001) to randomly assign for each node a block label which is indexed from 1 to K, where $K \in \{5, 10, 20\}$. We then set $P(a_{ij} = 1) = 0.3N^{-0.3}$ if i and j are in the same block, and $P(a_{ij} = 1) = 0.3N^{-1}$. This indicates that the nodes within the same block have higher probability to connect than between blocks.

Example 3. (Power-law Distribution Network) According to Barabási and Albert (1999), it is a common phenomenon that the majority nodes in the network have small links, but a small amount of nodes have large number of links. The degrees of nodes could then be characterized by the power-law distribution. To generate the network structure in the spirit of this phenomenon, we simulate A as follows according to Clauset et al. (2009). For each node, we generate the in-degree $d_i = \sum_j a_{ji}$ according to the discrete power-law distribution as $P(d_i = k) = ck^{-\alpha}$, where c is a normalizing constant and the exponent parameter α is set to be $\alpha = 2.5$ by Clauset et al. (2009). Lastly, for the ith node, we randomly select d_i nodes as its followers.

5.2. Performance Measurements and Simulation Results

We consider different network sizes (i.e., $N = 100, 500, 1000$) and let $T = N/10$. we have considered the case of $N > T$, and the results are not significantly different. Moreover, for each example, the numerical performance is evaluated at $\tau = 0.1, 0.2, \cdots, 0.9$. The experiment is randomly replicated $R = 1000$ times for a reliable evaluation. Specifically, use $\hat{\theta}(\tau) = \{\hat{\beta}_0^{(m)}(\tau), \hat{\beta}_1^{(m)}(\tau), \hat{\beta}_2^{(m)}(\tau), \hat{\gamma}^{(m)\top}(\tau)\}^\top$ be the estimator from the mth replication. To evaluate the finite sample performance, the following measures are considered. First, for the given parameter $\beta_j(\tau)$ $(0 \le j \le 2)$, the root mean square error (RMSE) is calculated by $\text{RMSE}_j(\tau) = \{R^{-1} \sum_r (\hat{\beta}_j^{(r)}(\tau) - \beta_j(\tau))^2\}^{1/2}$. Besides, for the nodal effect function γ, the RMSE is given by $\text{RMSE}_\gamma(\tau) = \{(5R)^{-1}\sum_r \|\hat{\gamma}^{(r)}(\tau) - \gamma(\tau)\|^2\}^{1/2}$. Next, for each $\beta_j(\tau)$, a 95% confidence interval is constructed as $\text{CI}_j^{(r)}(\tau) = (\hat{\beta}_j^{(r)}(\tau) - z_{0.975}\widehat{\text{SE}}_j^{(r)}(\tau), \hat{\beta}_j^{(r)}(\tau) + z_{0.975}\widehat{\text{SE}}_j^{(r)}(\tau))$, where $\widehat{\text{SE}}_j^{(r)}(\tau)$ is the jth diagonal element of $(NT)^{-1}\tau(1-\tau)\widehat{\Sigma}_\theta$, and z_α is the αth quantile of the standard normal distribution. Then, the coverage probability (CP) can be computed as $\text{CP}_j(\tau) = R^{-1}\sum_{m=1}^{R} I\{\beta_j(\tau) \in \text{CI}_j^{(r)}(\tau)\}$, where $I(\cdot)$ is the indicator function. Lastly, the network density (ND) is given by $\{N(N-1)\}^{-1}\sum_{i_1,i_2} a_{i_1 i_2}$.

The three types of network structures are visualized in Figure 1. Besides, the detailed results of the three simulation examples are given in Table 1 to 3. It can be found for a fixed τ, the RMSE is decreased as N and T increased. For example, the RMSE of $\hat{\beta}_1(\tau)$ drops from 11.22×10^{-2} to 4.90×10^{-2} at $\tau = 0.1$ as N is increased from 100 to 500 in Example 1 for the t- distribution. It can also be noted that the RMSE for t-distribution of same network size N is slightly larger than standard normal distribution. At the same time, the network is becoming sparser as N increases (e.g. ND drops from 2.4% to 0.2% for the power-law distribution network from $N = 100$ to 1000). Moreover, it can be concluded the computed coverage probabilities for $\beta_j(\tau)$s are stable at the nomial level 95%, which corroborates with the theoretical results. Lastly, we plot the estimated $\hat{\beta}_j(\tau)$ with the 95% confidence interval against τ in Figure 2. A monotonic increasing pattern can be detected.

5.3. Financial Contagion and Shared Ownership

In this study, we focus on studying the financial risk contagion mechanism accounting for the common shared ownership information using NQAR model. Specifically, we apply our NQAR model to the Chinese Stock Market in 2013. The dataset consists of 2,442 stocks from the Chinese A share market, which are traded in Shanghai Stock

Exchange and Shenzhen Stock Exchange. Specifically, the corresponding response Y_{it} is the weekly absolute return volatility. The time series of averaged stock volatility is plotted in the left panel of Figure 3, where a relatively higher volatility level can be captured during May and July. To construct the network structure, the top ten shareholders' information for each stock are collected. For the ith and jth stock, $a_{ij} = 1$ if they share at least one common shareholder, otherwise $a_{ij} = 0$. The network structures of stocks with top 100 market values are visualized in the right panel of Figure 3. The resulting network density is 3.9%.

Besides the shared ownership information, the firm specific variables are also taken into consideration. Motivated by Fama and French (2015), we consider the following $K = 6$ covariates to represent stocks' fundamentals. They are, SIZE (measured by the logarithm of market value), BM (book to market ratio), PR (increased profit ratio compared to the last year), AR (increased asset ratio compared to the last year), LEV (leverage ratio), and Cash (cash flow of the firm). Lastly, all covariates are re-scaled within the interval $[0, 1]$.

We then launch the network analysis using the NQAR model. Particularly, the estimation results of our NQAR model are given in Table 4 for $\tau = 0.05$, 0.50, and 0.95 respectively. It is noteworthy that the stocks have stronger network effect and momentum effect in the upper tail (i.e., $\tau = 0.95$) than the median and lower tail case (i.e., $\tau = 0.5$ and 0.05). This indicates that stocks tend to have higher correlation through the network when higher volatility level is exposed in the market. Besides, the size (i.e., CAP), the book to market ratio (BM), and the leverage ratios (LEV) are shown to have a negative correlation with the conditional quantile level of the volatility at $\tau = 0.95$ and $\tau = 0.5$. However, this phenomena does not appear in the lower quantile case. In addition, stocks with higher liquidity (i.e., cash flow), profit ratio (PR), and asset increasing ratio (AR) tend to have lower volatility in the median case.

Lastly, we conduct an impulse analysis in Section 3. Particularly, the influential power can be calculated by $\hat{\nu} = \{(1-\hat{b}_2)I_N - \hat{b}_1 W^\top\}^{-1}\mathbf{1}$, where $\hat{b}_1 = 10^{-1}\sum_{m=0}^{9}\hat{\beta}_1(0.05+0.1m)$ and $\hat{b}_2 = 10^{-1}\sum_{m=0}^{9}\hat{\beta}_2(0.05+0.1m)$ are the numerical approximations for b_1 and b_2. Specifically, the influential power is found to have a linear pattern with the weighted degrees in the left panel of Figure 4, and the histogram of the weighted degrees is given in the right panel of Figure 4. Then, the cross-sectional impulse analysis is conducted. Particularly, $\mathrm{IEI}_{l,\tau}$ ($\tau = 0.05, 0.5, 0.95$) is computed within 5 banks at $l = 1, \cdots, 10$, which is visualized in Figure 5. Note that the impulse direction is from column to row. The banks include, Bank of China (BOC), China Merchants Bank (CMB), Industrial and Commercial Bank of China (ICBC), Ping An Bank (PAB), and Shanghai Pudong Development Bank (SPDB). We observe significant asymmetric effects across different quantiles. It can be seen the IEI is higher in the upper tail ($\tau = 0.95$). Moreover, the impulse impacts between the BOC, CMB, and ICBC are much stronger than with the other two banks.

APPENDIX A

In Appendix A, we are going to prove the stationarity result (Theorem 2.1) in Section A.1. the proof of Theorem 2.2 are given in Section A.2, and the verification of (2.5) and (2.6) are given in Section A.3.

Appendix A.1: Proof of Theorem 2.1

First of all, by iteration, we obtain the solution of NQAR model (2.1) as

$$\mathbb{Y}_t = \sum_{l=0}^{L-1}\Pi_l\Gamma + \Pi_L\mathbb{Y}_{t-L} + \sum_{l}^{L-1}\Pi_l V_{t-l} = \sum_{l=0}^{\infty}\Pi_l\Gamma + \sum_{l=0}^{\infty}\Pi_l V_{t-l}, \tag{A.1}$$

where $\Pi_l \stackrel{\text{def}}{=} G_t G_{t-1} \cdots G_{t-l+1}$ for $l > 0$ and $\Pi_0 = 1$. We then prove Theorem 2.1 in two steps. In the first step, we prove the covariance stationarity of the solution (A.1). Next, we prove the uniqueness of the stationary solution (A.1).

Step 1. (Proof of Covariance Stationarity)

In this step, we show the covariance stationarity by calculating $\mathsf{E}(\mathbb{Y}_t)$ and $\mathsf{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-h})$ respectively. Denote $\lambda_i(M)$ as the ith eigenvalue of any arbitrary matrix $M \in \mathbb{R}^{N\times N}$ such that $|\lambda_1(M)| > |\lambda_2(M)| > \cdots > |\lambda_N(M)|$. Recall that $\mathsf{E}(G_t) = G = b_1 W + b_2 I$. We firstly verify that $\mathsf{E}(\mathbb{Y}_t) = \mu_Y$ for $1 \le t \le T$. To this end, note that $|\lambda_1(W)| = 1$ by Banerjee et al. (2014), we have

$$|\lambda_1(G)| \le |b_1||\lambda_1(W)| + |b_2| < 1 \tag{A.2}$$

due to the stationarity condition in Theorem 2.1. Then it could be computed that $\mathsf{E}(\mathbb{Y}_t) = \sum_{l=0}^{\infty} G^l \Gamma = (I-G)^{-1}\Gamma$ due to the independence of G_ts over t, and $\mathsf{E}(V_{t-l}) = 0$ for $l \ge 0$. Recall that we have $\Gamma = c_0 \mathbf{1}_N$. Then it is straightforward to have $\mu_Y = c_1^{-1} c_0 \mathbf{1}_N$, where $c_1 = (1 - b_1 - b_2)^{-1}$ is defined in Theorem 2.1.

We next calculate the covariance of $\mathbb{Y}_t$. Specifically, it can be expressed as

$$\begin{aligned}\mathsf{Cov}(\mathbb{Y}_t) = &\mathsf{Cov}(\sum_{l=0}^{\infty} \Pi_l \Gamma) + \mathsf{Cov}(\sum_{l_1=0}^{\infty} \Pi_{l_1}\Gamma, \sum_{l_2=0}^{\infty} \Pi_{l_2} V_{t-l_2}) \\ &+ \mathsf{Cov}(\sum_{l_2=0}^{\infty} \Pi_{l_2} V_{t-l_2}, \sum_{l_1=0}^{\infty} \Pi_{l_1}\Gamma) + \mathsf{Cov}(\sum_{l=0}^{\infty} \Pi_l V_{t-l}).\end{aligned} \tag{A.3}$$

Recall that $G^* = \mathsf{E}(G_t \otimes G_t) = \mathsf{E}\{\mathbf{B}_1(U_t) \otimes \mathbf{B}_1(U_t)\}(W \otimes W) + \mathsf{E}\{\mathbf{B}_1(U_t) \otimes \mathbf{B}_2(U_t)\}(W \otimes I) + \mathsf{E}\{\mathbf{B}_2(U_t) \otimes \mathbf{B}_1(U_t)\}(I \otimes W) + \mathsf{E}\{\mathbf{B}_2(U_t) \otimes \mathbf{B}_2(U_t)\}(I \otimes I)$, $\widetilde{b}_1 = \{\mathsf{E}\,\beta_1^2(U_{it})\}^{1/2}$, and $\widetilde{b}_2 = \{\mathsf{E}\,\beta_2^2(U_{it})\}^{1/2}$. Then we have

$$|\lambda_1(G^*)| \le \widetilde{b}_1^2|\lambda_1(W)|^2 + 2\widetilde{b}_1\widetilde{b}_2|\lambda_1(W)| + \widetilde{b}_2^2 \le (\widetilde{b}_1 + \widetilde{b}_2)^2 < 1, \tag{A.4}$$

by the fact that $|\lambda_1(W)| < 1$ and the stationarity condition in Theorem 2.1. Note the matrix G^* can be represented in Jordan canonical form as $P\Lambda P^{-1}$, where Λ is a matrix of the Jordan block diagonal form with diagonal elements being $\lambda_i(G^*)$ $(1 \le i \le N)$ and P is an invertible matrix. Then by (A.4), $(G^*)^l$ converges to zero at a geometric rate as $l \to \infty$ and therefore we have

$$\sum_{l=0}^{\infty}(G^*)^l = (I - G^*)^{-1}. \tag{A.5}$$

Similarly, by (A.4) we have $\sum_{l=0}^{\infty} G^l = (I - G)^{-1}$. We then calculate the terms of $\mathsf{Cov}(\mathbb{Y}_t)$ in (A.3) one by one.

For the first term it can be calculated $\mathsf{Cov}(\sum_{l=0}^{\infty}\Pi_l\Gamma) = \mathsf{E}\{(\sum_{l_1}^{\infty}\Pi_{l_1}\Gamma)(\sum_{l_2}^{\infty}\Gamma^\top\Pi_{l_2}^\top)\} - \mu_Y\mu_Y^\top = \sum_{l_1,l_2=0}^{\infty}\mathsf{E}(\Pi_{l_1}\Gamma\Gamma^\top\Pi_{l_2}^\top) - \mu_Y\mu_Y^\top$. Firstly we have $\mathrm{vec}(\Pi_{l_1}\Gamma\Gamma^\top\Pi_{l_2}^\top) = (\Pi_{l_2} \otimes \Pi_{l_1})\mathrm{vec}(\Gamma\Gamma^\top)$. Without loss of generality, we assume $l_1 \ge l_2$. Then it can be obtained $\mathsf{E}(\Pi_{l_2} \otimes \Pi_{l_1}) = (G^*)^{l_2}(I_N \otimes G)^{l_1 - l_2}$ and

$$\mathsf{E}(\Pi_{l_2} \otimes \Pi_{l_1})\mathrm{vec}(\Gamma\Gamma^\top) = (G^*)^{l_2}(b_1 + b_2)^{l_1 - l_2}c_0^2\mathbf{1}_{N^2} \tag{A.6}$$

due to that $\mathrm{vec}(\Gamma\Gamma^\top) = c_0^2\mathbf{1}_{N^2}$, $(I_N \otimes G)\mathbf{1} = (b_1 + b_2)\mathbf{1}$. Therefore, by (A.4), (A.5), and (A.6) we have $\sum_{l_1,l_2=0}^{\infty}\mathsf{E}\{\mathrm{vec}(\Pi_{l_1}\Gamma\Gamma^\top\Pi_{l_2}^\top)\} = \{\sum_{l_2=0}^{\infty}(G^*)^{l_2}\sum_{l_1>l_2}(b_1 + b_2)^{l_1 - l_2} + \sum_{l_1=0}^{\infty}(G^*)^{l_1}\sum_{l_1 \le l_2}(b_1 + b_2)^{l_2 - l_1}\}c_0^2\mathbf{1} = M_1\mathbf{1}_{N^2}$, where $M_1 = c_1^{-1}c_0^2(1 + b_1 + b_2)(I - G^*)^{-1}$. As a result, we have $\mathrm{vec}\{\mathsf{Cov}(\sum_{l=0}^{\infty}\Pi_l\Gamma)\} = M_1\mathbf{1}_{N^2} - c_1^{-2}c_0^2\mathbf{1}_{N^2}$.

Next, for the second term, we have $\mathsf{Cov}(\sum_{l_1=0}^{\infty}\Pi_{l_1}\Gamma, \sum_{l_2=0}^{\infty}\Pi_{l_2}V_{t-l_2}) = \sum_{l_1,l_2=0}^{\infty}\mathsf{E}(\Pi_{l_1}\Gamma V_{t-l_2}^\top\Pi_{l_2}^\top)$, due to that $\mathsf{E}(\sum_{l_2=0}^{\infty}\Pi_{l_2}V_{t-l_2}) = 0$. It is straightforward to verify that for $l_2 \le l_1$, we have $\mathsf{E}(\Pi_{l_1}V_{t-l_1}\Gamma^\top\Pi_{l_2}^\top) = 0$. Therefore, by (A.4) and (A.5), one could verify $\sum_{l_1,l_2=0}^{\infty}\mathsf{E}\{\mathrm{vec}(\Pi_{l_1}V_{t-l_1}\Gamma^\top\Pi_{l_2}^\top)\} = \sum_{l_1=0}^{\infty}\sum_{l_2=l_1+1}^{\infty}(G^*)^{l_1}\mathsf{E}\{(G_{t-l_1}\otimes I)(G\otimes I)^{l_2-l_1-1}(I\otimes V_{t-l_1})\}\Gamma = \sum_{l_1=0}^{\infty}\sum_{l_2=l_1+1}^{\infty}(G^*)^{l_1}(b_1+b_2)^{l_2-l_1-1}\mathrm{vec}(\Sigma_{bv}) = c_1^{-1}c_0(I-G^*)^{-1}\mathrm{vec}(\Sigma_{bv})$. Similarly, one could verify that the third term $\mathsf{Cov}(\sum_{l_1=0}^{\infty}\Pi_{l_1}V_{t-l_1}, \sum_{l_2=0}^{\infty}\Pi_{l_2}\Gamma)$ is also equivalent to $c_1^{-1}c_0(I-G^*)^{-1}\mathrm{vec}(\Sigma_{bv})$.

For the last term, we have $\mathsf{Cov}(\sum_{l=0}^{\infty}\Pi_l V_{t-l}) = \sum_{l=0}^{\infty}\mathsf{E}(\Pi_l V_{t-l}V_{t-l}^\top\Pi_l^\top)$ due to that $\mathsf{E}(\Pi_{l_1}V_{t-l_1}V_{t-l_2}^\top\Pi_{l_2}^\top) = 0$ for any $l_1 \ne l_2$. Then, note that $\mathsf{E}\{\mathrm{vec}(\Pi_l V_{t-l}V_{t-l}^\top\Pi_l^\top)\} = \mathsf{E}\{(\Pi_l\otimes\Pi_l)\mathrm{vec}(V_{t-l}V_{t-l}^\top)\} = (G^*)^l\mathrm{vec}(\Sigma_V)$. Then by (A.5) we have $\mathsf{Cov}(\sum_{l=0}^{\infty}\Pi_l V_{t-l}) = \sum_{l=0}^{\infty}(G^*)^l\mathrm{vec}(\Sigma_V) = \sum_{l=0}^{\infty}P\Lambda^l P^{-1}\mathrm{vec}(\Sigma_V) = P(I-\Lambda)^{-1}P^{-1}\mathrm{vec}(\Sigma_V) = (I-G^*)^{-1}\mathrm{vec}(\Sigma_V)$. Consequently, by (A.4) one obtains that $\mathrm{vec}\{\mathsf{Cov}(\sum_{l=0}^{\infty}\Pi_l V_{t-l})\} = (I-G^*)^{-1}\mathrm{vec}(\Sigma_V)$.

From the above, we have $\mathrm{vec}(\Sigma_Y)$ takes the form $(I-G^*)^{-1}\mathrm{vec}(\Sigma_V) + 2c_1^{-1}c_0(1-G^*)^{-1}\mathrm{vec}(\Sigma_{bv}) + (M_1 - c_1^{-2}c_0^2)\mathbf{1}_{N^2}$. To prove the covariance stationary, it suffices to prove that $\mathsf{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-h})$ only depends on h. As we can see that for $h \ge 1$, $\mathsf{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-h}) = \mathsf{E}(Y_tY_{t-h}^\top) - \mathsf{E}(Y_t)\,\mathsf{E}(Y_{t-h})^\top = \mathsf{E}\{\mathsf{E}(Y_t|\mathcal{F}_{t-h})Y_{t-h}\} - \mu_y\mu_y^\top$. Further one can obtain that $\mathsf{E}(Y_t|\mathcal{F}_{t-h}) = \sum_{l=0}^{h-1}G^l\Gamma + G^hY_{t-h}$. So it is straightforward to conclude $\mathsf{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-h}) = G^h\Sigma_Y$, which is only related to h. This completes the proof of Step 1.

Step 2. (Uniqueness of the Solution)

Assume that $\mathbb{Y}_t^*$ is another covariance stationary solution to the NQAR model. Define $|M|_a = (|m_{ij}|) \in \mathbb{R}^{m\times n}$ for any arbitrary matrix $M \in \mathbb{R}^{m\times n}$. In addition, for any two arbitrary matrices $M_1 = (m_{1,ij}) \in \mathbb{R}^{m\times n}$ and $M_2 = (m_{2,ij}) \in \mathbb{R}^{m\times n}$, define $M_1 \preceq M_2$ as $m_{1,ij} \le m_{2,ij}$ for $1 \le i \le m$ and $1 \le j \le n$. Then we know that $\mathsf{E}\,|\mathbb{Y}_t^*|_a \preceq C_1\mathbf{1}_N$ for some constant C_1. Similarly we have $\mathbb{Y}_t^* = \sum_{l=0}^{m-1}\Pi_l(\Gamma + V_{t-l}) + \Pi_m\mathbb{Y}_{t-m}^*$. To calculate the difference between $\mathbb{Y}_t$ and $\mathbb{Y}_t^*$, we have $\mathsf{E}\,|\mathbb{Y}_t - \mathbb{Y}_t^*|_a = \mathsf{E}\,|(\sum_{l=m}^{\infty}\Pi_l\Gamma + \sum_{l=m}^{\infty}\Pi_l V_{t-l}) - \Pi_m\mathbb{Y}_{t-m}^*|_a \preceq C_2(\sum_{l=m}^{\infty}\mathsf{E}\,|\Pi_l|_a + \mathsf{E}\,|\Pi_m|_a)\mathbf{1}_N$, where $C_2 = \max\{C_1, c_0, \mathsf{E}\,|V_{it}|\}$. It can be verified that $\mathsf{E}\,|\Pi_l|_a \mathbf{1}_N = \mathsf{E}\,|\beta_1(U_{it})W + \beta_2(U_{it})I_N|_a^l \mathbf{1}_N \preceq (\widetilde{b}_1 + \widetilde{b}_2)^l \mathbf{1}_N$. Therefore we have $(\sum_{l=m}^{\infty} \mathsf{E}\,|\Pi_l|_a + \mathsf{E}\,|\Pi_m|_a)\mathbf{1}_N \preceq C_3(\widetilde{b}_1 + \widetilde{b}_2)^m \mathbf{1}_N$ for some positive constant C_3. As this holds for any m, we can then prove that $\mathbb{Y}_t = \mathbb{Y}_t^*$ by the stationary condition that $\widetilde{b}_1 + \widetilde{b}_2 < 1$ with probability 1. This completes the proof.

Appendix A.2: Proof of Theorem 2.2

In this subsection, we establish the asymptotic normality of $\mathbb{Y}_t$. Define $\widetilde{\mathbb{Y}}_t = \mathbb{Y}_t - \mu_Y = (\widetilde{Y}_{1t}, \cdots, \widetilde{Y}_{Nt})^\top \in \mathbb{R}^N$. We then adopt the dependent Lindeberg central limit theorem (theorem 2) in Bardet et al. (2008) on $(NT)^{-1/2}\widetilde{\mathbb{Y}}_t$. We verify the two conditions in the following two parts. Step 1 is concerning moments bounds, and Step 2 is regarding the time dependency.

Step 1. (Bounding Moments) First, it suffices to show that there exists $0 < \delta \le 1$ satisfying

$$S_T = (NT)^{-(2+\delta)/2} \mathbf{1}_N^\top \sum_{t=0}^{T} \mathsf{E}\,|\widetilde{\mathbb{Y}}_t|_a^{2+\delta} \to 0 \tag{A.7}$$

as $T \to \infty$, where $|Y|_a^{2+\delta}$ denotes $(|Y_1|^{2+\delta}, \cdots, |Y_p|^{2+\delta})^\top \in \mathbb{R}^p$ for a p-dimensional vector $Y = (Y_1, \cdots, Y_p)^\top \in \mathbb{R}^p$. Then, one can verify that $\mathsf{E}\,|\widetilde{\mathbb{Y}}_t|_a^{2+\delta} = \mathsf{E}\,|\sum_{l=0}^{\infty}(\Pi_l \Gamma + \Pi_l V_{t-l}) - \mu_Y|_a^{2+\delta}$. Further by the Jensen's inequality $S_T = (NT)^{-(2+\delta)/2}\mathbf{1}_N^\top \sum_{t=0}^{T} \mathsf{E}\,|\widetilde{\mathbb{Y}}_t|_a^{2+\delta} \le (NT)^{-(2+\delta)/2}\mathbf{1}_N^\top \sum_{t=0}^{T}\{\mathsf{E}\,|\sum_{l=0}^{\infty}\Pi_l V_{t-l}|_a^{2+\delta} + \mathsf{E}\,|\sum_l \Pi_l \Gamma|_a^{2+\delta} + \mathsf{E}\,|\mu_Y|_a^{2+\delta}\}(3^{2+\delta}/3)$. It is not hard to see that $(NT)^{-(2+\delta)/2}\mathbf{1}_N^\top \sum_{t=1}^{T} |\mu_Y|_a^{2+\delta} \to 0$. Let $\delta = 1$ and define $S_{Tv} = N^{-3/2}T^{-1/2}\mathbf{1}_N^\top \mathsf{E}\,|\sum_{l=0}^{\infty}\Pi_l V_{t-l}|_a^3$. Then it suffices to show $S_{Tv} \to 0$. It can be derived

$S_{Tv} \cdot = S_{Tv1} + S_{Tv2} + S_{Tv3} + S_{Tv4}$, where

$$S_{Tv1} = N^{-3/2}T^{-1/2}\mathbf{1}_N^\top \mathsf{E}\Big\{\sum_l |\Pi_l V_{t-l}|_a^3\Big\},$$

$$S_{Tv2} = 3N^{-3/2}T^{-1/2}\sum_{l_1}\sum_{l_2>l_1} \mathsf{E}\big\{|\Pi_{l_1}V_{t-l_1}|_a^2 \circ |\Pi_{l_2}V_{t-l_2}|_a\big\},$$

$$S_{Tv3} = 3N^{-3/2}T^{-1/2}\sum_{l_1}\sum_{l_2>l_1} \mathsf{E}\big\{|\Pi_{l_1}V_{t-l_1}|_a \circ |\Pi_{l_2}V_{t-l_2}|_a^2\big\},$$

$$S_{Tv4} = 6N^{-3/2}T^{-1/2}\sum_{l_1}\sum_{l_2>l_1}\sum_{l_3>l_2>l_1} \mathsf{E}\big\{|\Pi_{l_1}V_{t-l_1}|_a \circ |\Pi_{l_2}V_{t-l_2}|_a \circ |\Pi_{l_3}V_{t-l_3}|_a\big\},$$

and $\circ$ means point wise product. We then verify the terms $S_{Tvj} \to 0$ for $j = 1, \cdots, 4$ as follows.

Firstly we have $S_{Tv1} \le N^{-3/2}T^{-1/2}\sum_{l=0}^{\infty}\mathsf{E}\,|\Pi_l(V_{t-1})|_a^3 \preceq N^{-3/2}T^{-1/2}\sum_{l=0}^{\infty} C_3\mathsf{E}\,|(|\Pi_l|_a\mathbf{1}_N)|_a^3$, where $C_3 = \max_i \mathsf{E}\,|V_{it}|^3$ is finite by Theorem 2.2. Further, the above term is element-wisely bounded by $C_3N^{-3/2}T^{-1/2}\sum_{l=0}^{\infty} C_b^l\mathbf{1}_N$, where $C_b = \mathsf{E}(|\beta_1(U_{it})| + |\beta_2(U_{it})|)^3 < 1$ by Theorem 2.2. Consequently we have $S_{Tv1} \to 0$. Next we look at the second term in S_{Tv}. It can be firstly verified that $\mathsf{E}(3\sum_{l_1}\sum_{l_2>l_1}|\Pi_{l_1}V_{t-l_1}|_a^2 \circ |\Pi_{l_2}V_{t-l_2}|_a) \preceq 3C_v\sum_{l_1}\sum_{l_2>l_1}\mathsf{E}(|\Pi_{l_1}V_{t-l_1}|_a^2 \circ |\Pi_{l_2}\mathbf{1}_N|_a) \preceq 3C_v(\widetilde{b}_1 + \widetilde{b}_2)^{l_2-l_1}\sum_{l_1}\sum_{l_2>l_1}\mathsf{E}(|\Pi_{l_1}V_{t-l_1}|_a^2 \circ |\Pi_{l_1-1}\mathbf{1}_N|_a)$ due to the independence of Π_{l_1-1} and $\prod_{k=1}^{l_2-l_1} G_{t-l_1-k}$, and the inequality $\mathsf{E}(\prod_{k=1}^{l_2-l_1}|G_{t-l_1-k}|_a\mathbf{1}_N) \preceq (\widetilde{b}_1 + \widetilde{b}_2)^{l_2-l_1}\mathbf{1}_N$, where $C_v = \max_i \mathsf{E}(|V_{it}|)$. Further it can be derived that $\mathbf{1}_N^\top\mathsf{E}(|\Pi_{l_1}V_{t-l_1}|_a^2 \circ |\Pi_{l_1-1}\mathbf{1}_N|_a) =$

$$\mathsf{E}(|\Pi_{l_1}V_{t-l_1}|_a^{2\top}|\Pi_{l_1-1}\mathbf{1}_N|_a) \le C_{bv}\,\mathsf{E}(|\Pi_{l_1}\mathbf{1}_N|_a^{2\top}|\Pi_{l_1}\mathbf{1}_N|_a) \le NC_{bv}C_b^{l_1}, \tag{A.8}$$

where $C_{bv} = (\mathsf{E}\{V_{it}^4\})^{1/2}[\mathsf{E}\{(|\beta_1(U_{it})| + |\beta_2(U_{it})|)^2\}]^{1/2}$. As a result, we have $S_{Tv2} \to 0$. Then, by iteratively applying (A.8), one could obtain $S_{Tv3} \to 0$ and $S_{Tv4} \to 0$, where the details are omitted here. As a result, (A.7) can be obtained.

STEP 2. (TIME DEPENDENCY) Next, we verify conditions imposed on the dependency structure of $\widetilde{\mathbb{Y}}_t$. To this end, we show the definition of λ dependency as in Bardet et al. (2008).

DEFINITION 5.1. *(λ **dependency**) A process X_t in $\mathbf{R}^d$ is said to be λ dependent if there exists a sequence $\{\lambda_r\}$ such that $\lambda_r \to 0$ when $r \to \infty$ satisfying*

$$\Big|\mathsf{Cov}\{f(X_{m_1}, X_{m_2}, \cdots, X_{m_v}), g(X_{s_1}, X_{s_2}, \cdots, X_{s_u})\}\Big| \le (uvL_f L_g + vL_f + uL_g)\lambda_r,$$

for all $v, u \in N^* \times N^*$ (N^* denotes the natural number space), L_f and L_g as constants, where v, u are two integers corresponding to support of f and g respectively.

Next we prove that $T^{-1/2}\widetilde{\mathbb{Y}}_t$ is λ dependent with a satisfactory rate. For this propose, we rewrite the NQAR model to be $\widetilde{\mathbb{Y}}_t = G_t\widetilde{\mathbb{Y}}_{t-1} + V_t'$, and $V_t' = V_t + (G_t - G)\mu_Y$. Then we have $\widetilde{\mathbb{Y}}_t = \sum_l^\infty \Pi_l V_{t-l}'$. For convenience, we define $\widetilde{\mathbb{Y}}_t^L = \sum_l^L \Pi_l V_{t-l}'$ as the truncated form of $\widetilde{\mathbb{Y}}_t$.

First of all define $\Im_v = \{\widetilde{\mathbb{Y}}_{m1}, \widetilde{\mathbb{Y}}_{m2}, \cdots, \widetilde{\mathbb{Y}}_{m_v}\}$ and $\Im_u = \{\widetilde{\mathbb{Y}}_{s_1}, \widetilde{\mathbb{Y}}_{s_2}, \cdots, \widetilde{\mathbb{Y}}_{s_u}\}$. And $\Im_v^L = \{\widetilde{\mathbb{Y}}_{m_1}^L, \widetilde{\mathbb{Y}}_{m_2}^L, \cdots, \widetilde{\mathbb{Y}}_{m_v}^L\}$ and $\Im_u^L = \{\widetilde{\mathbb{Y}}_{s_1}^L, \widetilde{\mathbb{Y}}_{s_2}^L, \cdots, \widetilde{\mathbb{Y}}_{s_u}^L\}$. We than have $\mathsf{Cov}\left(f(\Im_v), g(\Im_u)\right) = \mathsf{Cov}\left(f(\Im_v) - f(\Im_v^L), g(\Im_u)\right) + \mathsf{Cov}\left(f(\Im_v^L), g(\Im_u) - g(\Im_u^L)\right) + \mathsf{Cov}\left(f(\Im_v^L), g(\Im_u^L)\right)$. Define $\tilde{f}(X) = f(X) - \mathsf{E}(f(X))$. Without loss of generality, we set $L = r - 1$. Then $\mathsf{Cov}(f(\Im_v^L), g(\Im_u^L)) = 0$, and $|\mathsf{Cov}(f(\Im_v), g(\Im_u))|$ can be bounded by $\|\widetilde{g}\|_\infty \mathsf{E}\,|\widetilde{f}(\Im_v) - \widetilde{f}(\Im_v^L)| + \|\widetilde{f}\|_\infty \mathsf{E}\,|\widetilde{g}(\Im_u) - \widetilde{g}(\Im_u^L)| \le c(vL_f + uL_g)\mathbf{1}_N^\top \mathsf{E}\,|\widetilde{\mathbb{Y}}_{ms} - \widetilde{\mathbb{Y}}_{ms}^L|_a$, where $ms = m_1 \wedge m_2$, c is a constant and $\|\cdot\|_\infty$ is the uniform norm of a function, which takes the supremum of the absolute value of a function on its support. Then it can be verified that $\mathsf{E}\,|\widetilde{\mathbb{Y}}_{ms} - \widetilde{\mathbb{Y}}_{ms}^L|_a = \mathsf{E}\,|\sum_{l=L+1}^\infty \Pi_l V_{t-l}'|_a \le \sum_{l=L+1}^\infty C_{v'} \mathsf{E}\,|\Pi_l \mathbf{1}_N|_a \preceq \sum_{l=L+1}^\infty C_{v'}[\mathsf{E}\{|\beta_1(U_{it})| + |\beta_2(U_{it})|\}]^l \mathbf{1}_N \le C_{v'}(\widetilde{b}_1 + \widetilde{b}_2)^{L+1}(1 - \widetilde{b}_1 - \widetilde{b}_2)^{-1}$, where $C_{v'} = \mathsf{E}(|V_{it}|) + 2(\widetilde{b}_1 + \widetilde{b}_2)$. As a result, it can be concluded $\mathsf{Cov}\left(f(\Im_v), g(\Im_u)\right) \to 0$ as $r \to \infty$. This completes the proof.

Appendix A.3: Verification of (2.5) and (2.6)

Assume $\widetilde{b}_1 = |\int_0^1 \beta_1(u)^2du|^{1/2} = o(1)$. Recall $\widetilde{b}_{22} = \int_0^1 \beta_2(u)^2du$, $\widetilde{b}_{12} = \int_0^1 \beta_1(u)\beta_2(u)du$. By the stationary condition we have $\widetilde{b}_{22} < 1$, then by the Cauchy's inequality we have $|\widetilde{b}_{12}| \leq \widetilde{b}_1|\widetilde{b}_{22}|^{1/2} = o(1)$. Recall that $\mathrm{vec}(\Sigma_Y) = S_1 + S_2 + S_3$, where $S_1 = M_1\mathbf{1}_{N^2} - c_1^{-2}c_0^2\mathbf{1}_{N^2}$ ($M_1 = c_1^{-1}c_0^2(1+b_1+b_2)(I-G^*)^{-1}$), $S_2 = 2c_1^{-1}c_0(1-G^*)^{-1}\mathrm{vec}(\Sigma_{bv})$, and $S_3 = (I - G^*)^{-1}\mathrm{vec}(\Sigma_V)$. Next, we approximate Σ_Y by neglecting higher order terms of $b_1, \widetilde{b}_{12}, \widetilde{b}_1$. To this end, we first approximate $(I - G^*)^{-1}$ and c_1^{-1} as follows

$$(I - G^*)^{-1} \approx (I - \widetilde{B}_{22})^{-1}(I + M_{12}), \tag{A.9}$$

$$c_1^{-1} \approx (1 - b_2)^{-1}\big\{1 + (1 - b_2)^{-1}b_1\big\}, \tag{A.10}$$

$$c_1^{-2} \approx (1 - b_2)^{-2}\big\{1 + 2(1 - b_2)^{-1}b_1\big\}, \tag{A.11}$$

where $M_{12} = (I - \widetilde{B}_{22})^{-1}\{\widetilde{B}_{12}(W \otimes I) + \widetilde{B}_{21}(I \otimes W)\}$, $\widetilde{B}_{22} = E\{B_2(U_t) \otimes B_2(U_t)\}$, $\widetilde{B}_{12} = E\{B_1(U_t) \otimes B_2(U_t)\}$, and $\widetilde{B}_{21} = E\{B_2(U_t) \otimes B_1(U_t)\}$.

Recall that $\widetilde{b}_{01} = \mathsf{E}\{\beta_1(U_{it})(\beta_0(U_{it}) - b_0)\}$ and $\widetilde{b}_{02} = \mathsf{E}\{\beta_2(U_{it})(\beta_0(U_{it}) - b_0)\}$. Then, by (A.9)-(A.11) one could verify that $S_1 \approx (I - \widetilde{B}_{22})^{-1}\{(1 + 2b_1 - b_2^2)I \otimes I + (1 - b_2^2)(I - \widetilde{B}_{22})^{-1}(\widetilde{B}_{12} + \widetilde{B}_{21})\}(1 - b_2)^{-2}c_0^2\mathbf{1}_{N^2} - \{1 + 2(1 - b_2)^{-1}b_1\}(1 - b_2)^{-2}c_0^2\mathbf{1}_{N^2}$, $S_2 \approx 2(1 - b_2)^{-1}(I - \widetilde{B}_{22})^{-1}\{\widetilde{b}_{02}I + (1 - b_2)^{-1}b_1\widetilde{b}_{02}I + \widetilde{b}_{02}M_{12} + \widetilde{b}_{01}I\}c_0\mathrm{vec}(I_N)$, and $S_3 \approx (I - \widetilde{B}_{22})^{-1}(I + M_{12})\mathrm{vec}(\Sigma_V)$. Let $S_j = \mathrm{vec}(\Sigma_j)$ for $j = 1, 2, 3$ and $\Sigma_j = (\Sigma_{j,kl}) \in \mathbb{R}^{N\times N}$. Specifically, one can verify for the diagonal elements that $\Sigma_{1,ii} \approx [(1 - \widetilde{b}_{22})^{-1}\{1 - b_2^2 + 2b_1 + 2(1 - \widetilde{b}_{22})^{-1}(1 - b_2^2)\widetilde{b}_{12}\} - (1 - b_2)^{-1}(1 - b_2 + 2b_1)](1 - b_2)^{-2}c_0^2$, $\Sigma_{2,ii} \approx 2(1 - \widetilde{b}_{22})^{-1}(1 - b_2)^{-2}\{\sigma_{bv}(1 - b_2) + b_1\widetilde{b}_{02}\}c_0$ ($\sigma_{bv} = \widetilde{b}_{01} + \widetilde{b}_{02}$), $\Sigma_{3,ii} \approx (1 - \widetilde{b}_{22})^{-1}\sigma_V^2$. Similarly, we have $\Sigma_{1,i_1i_2} \approx \{(1 - b_2)^{-2}(1 - b_2^2)^{-2}(1 - b_2^2 + 2b_1 + 2b_1b_2) - (1 - b_2)^{-3}(1 - b_2 + 2b_1)\}c_0^2$, $\Sigma_{2,i_1i_2} \approx 2(1 - b_2^2)^{-2}(1 - b_2)^{-1}c_0b_1b_2\widetilde{b}_{02}(w_{i_1i_2} + w_{i_2i_1})$, $\Sigma_{3,i_1i_2} \approx (1 - b_2^2)^{-2}b_1b_2(w_{i_1i_2} + w_{i_2i_1})\sigma_V^2$ for $i_1 \neq i_2$, where $w_{i_1i_2} = n_{i_1}^{-1}a_{i_1i_2}$ is the (i, j)th element of W. This leads to the desired results.

APPENDIX B

In Appendix B, we give the proof of the asymptotic properties in the estimation part. Specifically, Theorem 4.1 and Theorem 4.2 are going to be proved in Section B.2 respectively.

Appendix B.1: A Useful Lemma

In this section, we give the proof of a useful lemma, which is employed as tools in later proof of asymptotic properties.

LEMMA 5.1. *Assume $c_\beta < 1$ and (C1)–(C3), where c_β is defined in (C1). Let $U = (U_1, \cdots, U_N)^\top \in \mathbb{R}^N$ and $V = (V_1, \cdots, V_N)^\top \in \mathbb{R}^N$, where U_i and V_i are identically distributed respectively for $1 \leq i \leq N$, and independent with Π_l. Assume $\mathsf{E}(U_i^4)^{1/4} \leq \nu_u$, $\mathsf{E}(V_i^4)^{1/4} \leq \nu_v$, $\mathsf{Cov}(U_i, V_i) \neq 0$, and $\mathsf{Cov}(U_i, U_j) = 0$ for $i \neq j$. Define $\mathbb{G} = \|\beta_1\|_4 W + \|\beta_2\|_4 I \in \mathbb{R}^{N \times N}$. Then the following results hold.*
(a) For any integer $l_1, l_2, l_3, l_4 > 0$ we have

$$\mathsf{E}(|\Pi_{l_1}^\top \Pi_{l_2} \Pi_{l_3}^\top \Pi_{l_4}|_a) \preccurlyeq |\mathbb{G}^{l_1 \top} \mathbb{G}^{l_2} \mathbb{G}^{l_3 \top} \mathbb{G}^{l_4}|_a. \tag{B.1}$$

(b) There exists a finite integer $K > 0$, such that for any $l > 0$, we have

$$\mathbb{G}^l \mathbb{G}^{l \top} \preccurlyeq l^K c_\beta^{2l} \mathcal{M}, \tag{B.2}$$

where $\mathcal{M} = MM^\top$ with $M = c_m \mathbf{1} \pi^\top + \sum_{j=1}^K W^j$, $c_m > 1$ is a constant, and π is defined *in (C2.1). Denote $\mathcal{M}_{ij}$ as the (i,j)th element of $\mathcal{M}$. We then have*

$$N^{-2}\mathbf{1}^\top\mathcal{M}\mathbf{1} \to 0, \tag{B.3}$$

$$N^{-2}\, tr(\mathcal{M}^2) \to 0, \tag{B.4}$$

as $N \to \infty$.

(c) For any integer $l_1 \geq l_2$, it holds that

$$\mathsf{Var}(U^\top \Pi_{l_1}^\top \Pi_{l_2} V) \leq 8\nu_u^2\nu_v^2 c_\beta^{2(l_1+l_2)} l_1^{2K}\{\mathbf{1}^\top\mathcal{M}\mathbf{1} + tr(\mathcal{M}^2)\}. \tag{B.5}$$

(d) We have $\widehat{\Omega}_0 \to_p \Omega_0$ as $\min\{N,T\} \to \infty$.

Proof of (a). We first derive an inequality of $\mathsf{E}\,|\Pi_{l_1}^\top\Pi_{l_2}\Pi_{l_3}^\top\Pi_{l_4}|_a$ as

$$\mathsf{E}\,|\Pi_{l_1}^\top\Pi_{l_2}\Pi_{l_3}^\top\Pi_{l_4}|_a \preccurlyeq \mathsf{E}\Big\{\big(\prod_{l=0}^{l_1-1}|G_{t-l}|_a\big)^\top\big(\prod_{l=0}^{l_2-1}|G_{t-l}|_a\big)\big(\prod_{l=0}^{l_3-1}|G_{t-l}|_a\big)^\top\big(\prod_{l=0}^{l_4-1}|G_{t-l}|_a\big)\Big\}.$$

We first prove

$$\mathsf{E}(|G_t|_a^\top|G_t|_a M|G_t|_a^\top|G_t|_a) \preccurlyeq \mathbb{G}^\top\mathbb{G}\,\mathsf{E}(M)\mathbb{G}^\top\mathbb{G} \tag{B.6}$$

for any elementwisely positive stochastic matrix M, where $M = (m_{ij}) \in \mathbb{R}^{N\times N}$ is assumed to be independent with G_t. Let $\mathbb{W}_{11} = W^\top W$, $\mathbb{W}_{10} = W^\top$, $\mathbb{W}_{01} = W$, $\mathbb{W}_{00} = I$. Further denote $\mathbb{W}_{q_1q_2,i\cdot} \in \mathbb{R}^N$ as the ith row vector of $\mathbb{W}_{q_1q_2}$, where $q_1, q_2 = 0, 1$. Then one could verify the (i,j)th element of $\mathsf{E}(|G_t|_a^\top|G_t|_a M|G_t|_a^\top|G_t|_a)$ involves a sum of terms like $\mathsf{E}\{|\beta_1^{k_1}(U_{i_1t})\beta_2^{k_2}(U_{i_2t})\beta_1^{k_3}(U_{i_3t})\beta_2^{k_4}(U_{i_4t})|(\mathbb{W}_{q_1q_2}M\mathbb{W}_{q_3q_4})_{ij}\}$, where $q_1, q_2 = 0, 1$, k_1, k_2, k_3, k_4 are integers, $0 \leq k_i \leq 4$, $k_1+k_2+k_3+k_4 = 4$, $0 \leq i_1, i_2, i_3, i_4 \leq N$.

By Hölder's inequality, we have for all i_1, i_2, i_3, i_4, $\mathsf{E}\,|\beta_1^{k_1}(U_{i_1t})\beta_2^{k_2}(U_{i_2t})\beta_1^{k_3}(U_{i_3t})\beta_2^{k_4}(U_{i_4t})| \leq \|\beta_1\|_4^{k_1+k_3}\|\beta_2\|_4^{4-(k_1+k_3)}$. By applying the inequality one could obtain (B.6). Subsequently, (B.1) can be derived by recursively applying (B.6).

Proof of (b). Note we have $\|\beta_1\|_4+\|\beta_2\|_4 < 1$. Then (B.2) can be obtained by (5.1) of Lemma 2 (a) in the supplementary material of Zhu et al. (2016). For the completeness of the proof, we briefly repeat the step as below. Firstly, for any integer $l > 0$, we have $\mathbb{G}^l = (\|\beta_1\|_4 W + \|\beta_2\|_4 I)^l = \sum_{j=0}^{l} C_l^j \|\beta_1\|_4^j \|\beta_2\|_4^{l-j} W^j$, where $C_l^j = l!/\{j!(l-j)!\}$. Since W is an element-wise non-negative matrix, $|\mathbb{G}^l|_a \preccurlyeq \sum_{j=0}^{l} C_l^j \|\beta_1\|_4^j \|\beta_2\|_4^{l-j} W^j$. Then for $l > K$ we have $|\mathbb{G}^l|_a \preccurlyeq (\|\beta_1\|_4 + \|\beta_2\|_4)^l C\mathbf{1}\pi^\top + \sum_{j=0}^{K} C_l^j \|\beta_1\|_4^j \|\beta_2\|_4^{l-j} W^j$, , where this fact is due to $W^l \preccurlyeq C\mathbf{1}\pi^\top$. Further note that $\|\beta_1\|_4^j \|\beta_2\|_4^{n-j} < c_\beta^l$ $(0 \le j \le l)$, and $C_l^K \le l^K$. As a result, for $l > K$ we have,

$$|\mathbb{G}^l|_a \preccurlyeq l^K (\|\beta_1\|_4 + \|\beta_2\|_4)^l M, \tag{B.7}$$

where recall that $M = C\mathbf{1}\pi^\top + \sum_{j=0}^{K} W^j$. It is easy to verify that (B.7) also holds for $l = 1, \cdots, K-1$. Then we have $|\mathbb{G}^l(\mathbb{G}^\top)^l|_a \preccurlyeq l^{2K} c_\beta^{2l} M M^\top$ for any positive integer n. As a result, (B.2) can be proved.

Next (B.3) and (B.4) can be obtained by (5.11) and (5.12) of the supplementary material by Zhu et al. (2016) respectively.

Proof of (c). We first prove that (B.5) holds for $l_1 = l_2 = l$, then extend the results to $l_1 > l_2$. Let $l_1 = l_2 = l$, we have

$$\mathsf{Var}(U^\top \Pi_l^\top \Pi_l V) = \mathsf{Var}\{\mathsf{E}(U^\top \Pi_l^\top \Pi_l V | \Pi_l)\} + \mathsf{E}\{\mathsf{Var}(U^\top \Pi_l^\top \Pi_l V | \Pi_l)\}. \tag{B.8}$$

We then derive the upper bound for $\mathsf{E}\{\mathsf{Var}(U^\top \Pi_l^\top \Pi_l V)|\Pi_l\}$ and $\mathsf{Var}\{\mathsf{E}(U^\top \Pi_l^\top \Pi_l V)|\Pi_l\}$ in the following respectively.

Upper Bound for $\mathsf{E}\{\mathsf{Var}(U^\top \Pi_l^\top \Pi_l V)|\Pi_l\}$. One could first verify that $U^\top \Pi_l^\top \Pi_l V =$

$\mathrm{vec}(\Pi_l)^\top(I\otimes\Pi_l)\mathrm{vec}(VU^\top)$. Denote $\mathcal{V}=\mathrm{vec}(VU^\top)\in\mathbb{R}^N$. As a consequence, we have

$$\mathsf{Var}(U^\top\Pi_l^\top\Pi_l V|\Pi_l)=\mathrm{vec}(\Pi_l)^\top(I\otimes\Pi_l)\,\mathsf{Cov}(\mathcal{V})(I\otimes\Pi_l^\top)\mathrm{vec}(\Pi_l).$$

Further more, by the Cauchy's inequality for the $N\times N$ block matrices $\Sigma_{\mathcal{V},ii}$ and $\Sigma_{\mathcal{V},ij}$ the following bound can be attained,

$$\Sigma_{\mathcal{V},ii}=\mathsf{Cov}(V_iU)\le 2\nu_u^2\nu_v^2\mathbf{1}\mathbf{1}^\top, \tag{B.9}$$

$$\Sigma_{\mathcal{V},ij}=\mathsf{Cov}(V_iU,V_jU)\le 2\nu_u^2\nu_v^2(I+e_j\mathbf{1}^\top+\mathbf{1}e_i^\top) \tag{B.10}$$

for $i\neq j$, where $e_i\in\mathbb{R}^N$ is a vector with all elements to be 0 but only the ith element being 1. Denote $\Pi_{l,\cdot i}$ as the ith column vector of Π_l. Then we have $\mathrm{vec}(\Pi_l)^\top(I\otimes\Pi_l)\,\mathsf{Cov}(\mathcal{V})(I\otimes\Pi_l^\top)\mathrm{vec}(\Pi_l)=\sum_{i,j=1}^N\Pi_{l,\cdot i}^\top\Pi_l\Sigma_{\mathcal{V},ij}\Pi_l^\top\Pi_{l,\cdot j}=\sum_{i=1}^N\Pi_{l,\cdot i}^\top\Pi_l\Sigma_{\mathcal{V},ii}\Pi_l^\top\Pi_{l,\cdot i}+\sum_{i\neq j}\Pi_{l,\cdot i}^\top\Pi_l\Sigma_{\mathcal{V},ij}\Pi_l^\top\Pi_{l,\cdot j}\le 2\nu_u^2\nu_v^2\{3\,\mathrm{tr}(|\Pi_l^\top\Pi_l|_a\mathbf{1}\mathbf{1}^\top|\Pi_l^\top\Pi_l|_a)+\mathrm{tr}(|\Pi_l^\top\Pi_l|_a|\Pi_l^\top\Pi_l|_a)\}\le$
$6\nu_u^2\nu_v^2\mathbf{1}^\top|\Pi_l^\top|_a$
$|\Pi_l|_a|\Pi_l|_a^\top|\Pi_l|_a\mathbf{1}+2\nu_u^2\nu_v^2\,\mathrm{tr}(|\Pi_l^\top|_a|\Pi_l|_a|\Pi_l|_a^\top|\Pi_l|_a)$. By taking expectation on the right side we have

$$\mathsf{E}\left\{\mathsf{Var}(U^\top|\Pi_l^\top\Pi_l|_aV||\Pi_l|_a)\right\}\le 6\nu_u^2\nu_v^2c_\beta^{4l}l^{2K}\mathbf{1}^\top\mathcal{M}\mathbf{1}+2\nu_u^2\nu_v^2c_\beta^{4l}l^K\,\mathrm{tr}(\mathcal{M}^2). \tag{B.11}$$

Upper Bound for $\mathsf{Var}\{\mathsf{E}(U^\top\Pi_l^\top\Pi_lV|\Pi_l)\}$. It can be calculated that $\mathsf{E}(U^\top\Pi_l^\top\Pi_lV|\Pi_l)\le\nu_u\nu_v\,\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)$. Firstly we have $\mathsf{Var}\{\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)\}=\mathsf{E}[\mathsf{Var}\{\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)|\Pi_{l-1}\}]+$
$\mathsf{Var}[\mathsf{E}\{\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)$
$|\Pi_{l-1}\}]$. We first write $\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)=\sum_i G_{t-l+1,\cdot i}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_aG_{t-l+1,\cdot i}$. Therefore we have $\mathsf{Var}\{\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)|\Pi_{l-1}\}=\sum_i\mathsf{Var}(G_{t-l+1,\cdot i}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_aG_{t-l+1,\cdot i}|\Pi_{l-1})\le$

$2\sum_i(\mathbb{G}_{\cdot i}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G}_{\cdot i})^2$

$\leq 2c_\beta^2\sum_i \mathbb{G}_{\cdot i}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbf{1}\mathbf{1}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G}_{\cdot i} = 2c_\beta^2\mathbf{1}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G}\mathbb{G}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbf{1}$. Moreover, by similar proofs of (B.1), we have $\mathsf{E}(\mathbf{1}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G}\mathbb{G}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbf{1}) \leq l^K c_\beta^{4l-2}\mathbf{1}^\top\mathcal{M}\mathbf{1}$ by (B.1) and (B.3). Lastly, one could verify that $\mathsf{E}\{\mathrm{tr}(|\Pi_l|_a^\top|\Pi_l|_a)|\Pi_{l-1}\} \leq \sum_i \mathbb{G}_{\cdot i}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G}_{\cdot i} = \mathrm{tr}(\mathbb{G}^\top|\Pi_{l-1}|_a^\top|\Pi_{l-1}|_a\mathbb{G})$. By applying the deduction recursively, one should have $\mathsf{Var}\{\mathsf{E}(U^\top|\Pi_l|_a^\top|\Pi_l|_aV|\Pi_l)\} \leq 2\nu_u^2\nu_v^2c_\beta^{4l}\sum_{k=1}^l k^K\mathbf{1}^\top\mathcal{M}\mathbf{1}$. By combining the results of (B.11), we have

$$\mathsf{Var}(U^\top|\Pi_l|_a^\top|\Pi_l|_aV) \leq 2\nu_u^2\nu_v^2c_\beta^{4l}\Big\{3l^K\mathbf{1}^\top\mathcal{M}\mathbf{1} + \sum_{k=1}^{l}k^K\mathbf{1}^\top\mathcal{M}\mathbf{1} + l^{2K}\,\mathrm{tr}(\mathcal{M}^2)\Big\}. \quad \text{(B.12)}$$

Consequently we have (B.5) holds. For $l_1 > l_2$, it can be derived that $\mathsf{Var}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V) = \mathsf{Var}\{\mathsf{E}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\} + \mathsf{E}\{\mathsf{Var}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\}$. For $\mathsf{E}\{\mathsf{Var}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\}$, one can achieve a direct bound by (B.1) and (B.2) similar to the case of $l_1 = l_2$:
$\mathsf{E}\{\mathsf{Var}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\} \leq 2\nu_u^2\,\mathsf{E}(V^\top|\Pi_{l_2}|_a^\top|\Pi_{l_1}|_a|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV) \leq 2\nu_u^2\nu_v^2c_\beta^{2(l_1+l_2)}l_1^K\mathbf{1}^\top \mathcal{M}\mathbf{1}$.

For $\mathsf{Var}\{\mathsf{E}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\}$, we would like to bound by a recursive formula so that one can utilize the conclusion we achieved for the $l_1 = l_2$ case. $\mathsf{Var}\{\mathsf{E}(U^\top\Pi_{l_1}^\top\Pi_{l_2}V|\Pi_{l_1})\} \leq 2\nu_u^2\,\mathsf{Var}(\mathbf{1}^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV)$. So we have

$$\mathsf{Var}(U^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV) \leq 2\nu_u^2\nu_v^2c_\beta^{2(l_1+l_2)}l_1^K\mathbf{1}^\top\mathcal{M}\mathbf{1} + \nu_u^2\,\mathsf{Var}(\mathbf{1}^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV). \quad \text{(B.13)}$$

Note $\mathbf{1}^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV = \mathcal{B}_{t-l_1+1}^\top|\Pi_{l_1-1}|_a|\Pi_{l_2}|_aV$. Then by letting $U = \mathcal{B}_{t-l_1+1}$ one could obtain the result that $\mathsf{Var}(U^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV) \leq 2\nu_u^2\nu_v^2c_\beta^{2(l_1+l_2)}\{l_1^K + (l_1-1)^K\}\mathbf{1}^\top\mathcal{M}\mathbf{1} + \nu_u^2c_\beta^2\,\mathsf{Var}(\mathbf{1}^\top|\Pi_{l_1-1}|_a^\top|\Pi_{l_2}|_aV)$. By applying the same technique recursively, we have

$$\mathsf{Var}(U^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_aV) \leq 2\nu_u^2\nu_v^2c_\beta^{2(l_1+l_2)}\sum_{k=l_2+1}^{l_1}k^K\mathbf{1}^\top\mathcal{M}\mathbf{1}+\nu_u^2c_\beta^{2(l_2-l_1-1)}\,\mathsf{Var}(\mathcal{B}_{t-l_2}^\top|\Pi_{l_2}|_a^\top|\Pi_{l_2}|_aV).$$

By combining the results from (B.12), we have $\mathsf{Var}(U^\top|\Pi_{l_1}|_a^\top|\Pi_{l_2}|_a V) \le 2\nu_u^2\nu_v^2 c_\beta^{2(l_1+l_2)}\{(\sum_{k=1}^{l_1} k^K + 3l_2^K)\mathbf{1}^\top\mathcal{M}\mathbf{1} + l_2^{2K}\,\mathrm{tr}(\mathcal{M}^2)\} \le 8\nu_u^2\nu_v^2 c_\beta^{2(l_1+l_2)} l_1^{2K}\{\mathbf{1}^\top\mathcal{M}\mathbf{1}+\mathrm{tr}(\mathcal{M}^2)\}$, which proves (B.5).

Proof of (d): Write $\widehat{\Omega}_0 =$

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} X_{it}^\top X_{it} = \begin{pmatrix} 1 & \mathcal{S}_{12} & \mathcal{S}_{13} & \mathcal{S}_{14} \\ & \mathcal{S}_{22} & \mathcal{S}_{23} & \mathcal{S}_{24} \\ & & \mathcal{S}_{33} & \mathcal{S}_{34} \\ & & & \mathcal{S}_{44} \end{pmatrix},$$

where

$$\mathcal{S}_{12} = \frac{1}{N}\sum_{i=1}^{N} Z_i^\top, \quad \mathcal{S}_{13} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} w_i^\top \mathbb{Y}_t, \quad \mathcal{S}_{14} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} Y_{it},$$

$$\mathcal{S}_{22} = N^{-1}\sum_{i=1}^{N} Z_i Z_i^\top, \quad \mathcal{S}_{23} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} w_i^\top \mathbb{Y}_t Z_i, \quad \mathcal{S}_{24} = (NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} Y_{it} Z_i,$$

$$\mathcal{S}_{33} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} (w_i^\top \mathbb{Y}_t)^2, \quad \mathcal{S}_{34} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} w_i^\top \mathbb{Y}_t Y_{it}, \quad \mathcal{S}_{44} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} Y_{it}^2.$$

One can directly conclude that $\mathcal{S}_{12} \to_p \mathbf{0}_p^\top$ and $\mathcal{S}_{22} \to \Sigma_Z$ by the law of large numbers. Recall that $\kappa_1 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(\Sigma_Y)$, $\kappa_2 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(W\Sigma_Y)$, $\kappa_3 = \lim_{N\to\infty} N^{-1}\mathrm{tr}(W\Sigma_Y W^\top)$, and $\kappa_4 = \lim_{N\to\infty} N^{-1}\mathrm{tr}\{(I-G)^{-1}\}$, $\kappa_5 = \lim_{N\to\infty} N^{-1}\mathrm{tr}\{W(I-G)^{-1}\}$. $\Sigma_z = \mathsf{E}(Z_i Z_i^\top)$. $\mathcal{S}_{12} \to_p (0,0,\cdots,0)_{p\times 1}$, $\mathcal{S}_{13} \to_p c_b$, and $\mathcal{S}_{14} = N^{-1}\sum_t \mathbf{1}^\top \mathbb{Y}_{t-1} \to_p c_b$. We first list the component wise limit for each element in $\widehat{\Omega}_0$ in expectation, and we will verify the variance of these components in the next steps. $\mathcal{S}_{22} \to_p \Sigma_z$ and $\mathcal{S}_{23} = \frac{1}{NT}\sum_{t=1}^{T} Z^\top W \mathbb{Y}_{t-1} \to_p \frac{1}{N}\mathsf{E}\{Z^\top W(I-G)^{-1} Z\overline{\gamma}\} = \kappa_5 \Sigma_z \overline{\gamma}$, with $Z = (Z_1, Z_2, \cdots, Z_n)^\top$ and $\overline{\gamma} = [\int \gamma_l(u)du]$. $\mathcal{S}_{24} = \frac{1}{NT}\sum_{t=1}^{T} \mathbf{1}^\top \mathbb{Y}_{t-1} Z_i \to_p N^{-1}\mathsf{E}(Z^\top(I-G)^{-1}Z)\overline{\gamma} = \kappa_4 \Sigma_z \overline{\gamma}$.

$\mathcal{S}_{33} = \frac{1}{NT}\sum_t \sum_i (w_i^\top \mathbb{Y}_{t-1})^\top = \frac{1}{NT}\sum_{t=1}^T \mathbb{Y}_{t-1}^\top W^\top W \mathbb{Y}_{t-1} \to_p N^{-1}\operatorname{tr}\{W^\top W \Sigma_Y\} + c_b^2$. Finally $\mathcal{S}_{34} \to_p N^{-1}\operatorname{tr}\{W^\top \Sigma_Y\} + c_b^2$, $\mathcal{S}_{44} \to_p N^{-1}\operatorname{tr}\{\Sigma_Y\} + c_b^2$.

Some tedious steps are needed for verifying the variance of the aforementioned component. For the interest of space, we will only show one of the hardest part $\mathcal{S}_{44} \to_p \kappa_1 + c_b^2$ which involves the fourth moment of $\mathbb{Y}_t$. The proof contains two steps. In the first step, we prove that for any fixed t, $N^{-1}\sum_{i=1}^N X_{it}X_{it}^\top \to_p \Omega_0$ as $N \to \infty$. Next, we deal with the dependence cross over time (i.e., $1 \le t \le T$). Specifically, the near epoch dependence of Y_{it} and its functional forms are presented and consequently the desired law of large numbers results are established.

Step 1. Proof of $N^{-1}\sum_{i=1}^N Y_{it}^2 \to_p \kappa_1 + c_b^2$.

In this step, we prove $N^{-1}\sum_{i=1}^N Y_{it}^2 \to_p \kappa_1 + c_b^2$ as $N \to \infty$ for any fixed t under conditions (C1)–(C3). To this end, it suffices to show $N^{-1}\mathsf{E}(\mathbb{Y}_t^\top \mathbb{Y}_t) \to \kappa_1 + c_b^2$ and $N^{-2}\mathsf{Var}(\sum_t \mathbb{Y}_t^\top \mathbb{Y}_t) \to 0$ as $N \to \infty$. By condition (C3) we have $N^{-1}\mathsf{E}(\mathbb{Y}_t^\top \mathbb{Y}_t) = N^{-1}\{\operatorname{tr}(\Sigma_Y) + \mathsf{E}(\mathbb{Y}_t)^\top \mathsf{E}(\mathbb{Y}_t)\} \to \kappa_1 + c_b^2$. We then prove $N^{-2}\mathsf{Var}(\mathbb{Y}_t^\top \mathbb{Y}_t) \to 0$ as $N \to \infty$ in the following.

Recall that $\mathbb{Y}_t$ has the decomposition in the (A.1). Without loss of generality, assume $\Gamma = \mathbf{1}_N$. Then we have $\mathbb{Y}_t^\top \mathbb{Y}_t = \sum_{l_1,l_2=0}^\infty (\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2}\mathbf{1} + 2\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2} V_{t-l_2} + V_{t-l_1}^\top \Pi_{l_1}^\top \Pi_{l_2} V_{t-l_2})$. By the Cauchy's inequality, it suffices to show $N^{-2}\mathsf{Var}\left(\sum_{l_1,l_2=0}^\infty V_{t-l_1}^\top \Pi_{l_1}^\top \Pi_{l_2} V_{t-l_2}\right) \to 0$, $N^{-2}\mathsf{Var}\left(\sum_{l_1,l_2=0}^\infty \mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2} V_{t-l_2}\right) \to 0$, and $N^{-2}\mathsf{Var}\left(\sum_{l_1,l_2=0}^\infty \mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2}\mathbf{1}\right) \to 0$ as $N \to \infty$. Since their proofs are almost the same, we prove $N^{-2}\mathsf{Var}\left(\sum_{l_1,l_2=0}^\infty \mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2}\mathbf{1}\right) \to 0$ in the following for simplicity. To this end, first it can be shown $\mathsf{Var}(\sum_{l_1,l_2=0}^\infty \mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2}\mathbf{1}) =$

$\sum_{l_1 \neq l_2}^{\infty} \mathsf{Var}\left(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2} \mathbf{1}\right) + \sum_{l_1, l_2 = 0}^{\infty} \mathsf{Cov}\left(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_1} \mathbf{1}, \mathbf{1}^\top \Pi_{l_2}^\top \Pi_{l_2} \mathbf{1}\right)$. Then it suffices to show

$$N^{-2} \sum_{l_1 \neq l_2}^{\infty} \mathsf{Var}\left(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2} \mathbf{1}\right) \to 0, \tag{B.14}$$

$$N^{-2} \sum_{l_1, l_2 = 0}^{\infty} \mathsf{Cov}\left(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_1} \mathbf{1}, \mathbf{1}^\top \Pi_{l_2}^\top \Pi_{l_2} \mathbf{1}\right) \to 0 \tag{B.15}$$

$N \to \infty$. We then prove (B.14) and (B.15) separately as follows. Write $\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_2} \mathbf{1} = \mathcal{B}_{t-l_1+1}^\top \Pi_{l_1-1}^\top \Pi_{l_2-1} \mathcal{B}_{t-l_2+1}$, where $\mathcal{B}_{t-l_1+1} = \mathbf{B}_{1(t-l_1+1)} \mathbf{1}_N + \mathbf{B}_{2(t-l_1+1)} \mathbf{1}_N = (\beta_1(U_{i(t-l_1+1)})) + (\beta_2(U_{i(t-l_1+1)}))$. It can be calculated $\sum_{l_1 \neq l_2}^{\infty} \mathsf{Var}\left(\mathcal{B}_{t-l_1+1}^\top \Pi_{l_1-1}^\top \Pi_{l_2-1} \mathcal{B}_{t-l_2+1}\right) = 2 \sum_{l_2=0}^{\infty} \sum_{l_1 > l_2} \mathsf{Var}$ $\left(\mathcal{B}_{t-l_1+1}^\top \Pi_{l_1-1}^\top \Pi_{l_2-1} \mathcal{B}_{t-l_2+1}\right)$. By (B.5) we have $\mathsf{Var}\left(\mathcal{B}_{t-l_1+1}^\top \Pi_{l_1-1}^\top \Pi_{l_2-1} \mathcal{B}_{t-l_2+1}\right)$ $\leq 8 c_\beta^{2(l_1+l_2-2)} l_1^{2K} \{\mathbf{1}^\top \mathcal{M} \mathbf{1} + \mathrm{tr}(\mathcal{M}^2)\}$. Then we have (B.14) due to that $\sum_{l_2=0}^{\infty} \sum_{l_1 > l_2}$ $c_\beta^{2(l_1+l_2-2)} l_1^{2K} < \infty$ and $N^{-2}\{\mathbf{1}^\top \mathcal{M} \mathbf{1} + \mathrm{tr}(\mathcal{M}^2)\} \to 0$ by (B.3) and (B.4). For (B.15), it can be shown by Cauchy's inequality that $\mathsf{Cov}(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_1} \mathbf{1}, \mathbf{1}^\top \Pi_{l_2}^\top \Pi_{l_2} \mathbf{1}) \leq \mathsf{Var}(\mathbf{1}^\top \Pi_{l_1}^\top \Pi_{l_1} \mathbf{1})^{1/2}$ $\mathsf{Var}(\mathbf{1}^\top \Pi_{l_2}^\top \Pi_{l_2} \mathbf{1})^{1/2} \leq 8 c_\beta^{2(l_1+l_2)} l_1^K l_2^K \{\mathbf{1}^\top \mathcal{M} \mathbf{1} + \mathrm{tr}(\mathcal{M}^2)\}$ by (B.5). Then (B.15) holds since $\sum_{l_1, l_2} c_\beta^{2(l_1+l_2)} l_1^K l_2^K < \infty$ and $N^{-2}\{\mathbf{1}^\top \mathcal{M} \mathbf{1} + \mathrm{tr}(\mathcal{M}^2)\} \to 0$ as $N \to \infty$. This completes the proof.

Step 2. L_1 Near Epoch Dependence.

In this step, we further prove $N^{-1} \sum_i Y_{it}^2$ satisfies near epoch dependence cross $1 \leq t \leq T$. First we give the definition of L_1 near epoch dependence as below.

DEFINITION 5.2. *(L_1* ***near epoch dependence****) A triangular array U_{it} in $\mathbb{R}^1$ is said to be L_1 near epoch dependent (NED) if there exists constants c_{it} and a sequence $\{v_J, J \geq 1\}$ such that $v_J \to 0$ when $J \to \infty$ satisfying*

$$\mathsf{E}\left|(U_{it}) - \mathsf{E}(U_{it} | \mathcal{F}_{t-J}, \cdots, \mathcal{F}_t, \cdots, \mathcal{F}_{t+J})\right| \leq c_{it} v_J.$$

Given the definition, we firstly prove that Y_{it}s are L_1 NED by Andrews (1988). Next, according to Chapter 7 Lemma 1 of Gallant (2009), the smooth transformations of Y_{it}s (e.g., $N^{-1}\sum_i Y_{it}^2$) are also NED. Since Y_{it} has finite forth moment, then by Gallant (2009) we have $N^{-1}\sum_{i=1}^N Y_{it}^2$ is a uniformly integrable L_1 mixinggale. Consequently, according to Theorem 1 of Andrews (1988), we could have $(NT)^{-1}\sum_{t=1}^T\sum_{i=1}^N Y_{it}^2$ converge in probability as $N\to\infty$ and $T\to\infty$. We then prove Y_{it} is NED in the following.

Denote $\mathcal{F}_{t-J}^{t+J}=\{\mathcal{F}_{t-J},\cdots,\mathcal{F}_t,\cdots,\mathcal{F}_{t+J}\}$ and $\Pi_{t_1}^{t_2}=\prod_{t=t_1}^{t_2}G_t$. We then have the following inequality as

$$\begin{aligned}\mathsf{E}\Big\{e_i^\top\big|\mathbb{Y}_t-\mathsf{E}(\mathbb{Y}_t|\mathcal{F}_{t-J}^{t+J})\big|_a\Big\} &\leq \mathsf{E}\Big[e_i^\top\big\{\sum_{l=J+1}^{\infty}\Pi_l V_{t-l}+\sum_{l=J+1}^{\infty}\Pi_{J+1}(\Pi_{t-l-J}^{t-(J+1)}-G^{l-J-1})\Gamma\big\}\Big]\\ &\leq \sum_{l=J+1}^{\infty}(b_1^a+b_2^a)^l c_v+\sum_{l=J+1}^{\infty}2(b_1^a+b_2^a)^l c_0,\end{aligned}$$

where $c_v=\mathsf{E}|V_{it}|$. Let $v_J=(b_1^a+b_2^a)^{J+1}$ and $c_{it}=(1-b_1^a-b_2^a)^{-1}(2c_0+c_v)$. By condition (C1) we have $b_1^a+b_2^a<1$, thus Y_{it}s are L_1 NED according to Definition 5.2. This completes the proof of Step 2.

Appendix B.2: Proof of Theorem 4.1 and Theorem 4.2

Denote $V_{it\tau}=Y_{it}-X_{i(t-1)}^\top\theta(\tau)$ and $\hat{v}=\sqrt{NT}\big(\hat\theta(\tau)-\theta(\tau)\big)$. Then we have $\rho_\tau\big(Y_{it}-X_{i(t-1)}^\top\hat\theta(\tau)\big)=\rho_\tau\big(V_{it\tau}-(NT)^{-1/2}X_{i(t-1)}^\top\hat{v}\big)$, where $V_{it\tau}=Y_{it}-X_{i(t-1)}^\top\theta(\tau)$. Then the minimization of (4.1) is equivalent to minimizing

$$Z_{NT}(v)=\sum_{i=1}^N\sum_{t=1}^T\Big\{\rho_\tau\big(V_{it\tau}-(NT)^{-1/2}X_{i(t-1)}^\top v\big)-\rho_\tau(V_{it\tau})\Big\}.$$

One could verify that $\hat{v}=\arg\min_v Z_{NT}(v)$. The objective function $Z_{NT}(v)$ is a convex random function. Define $\psi_\tau(u)=\tau-I(u<0)$ and let $\nu_{it}=(NT)^{-1/2}v^\top X_{it}$, and one could further write $Z_{NT}(v)$ as $Z_{NT}(v) =$

$$-\sum_{i,t}\Big[(NT)^{-1/2}v^\top X_{i(t-1)}\psi_\tau(V_{it\tau})+\int_0^{\nu_{i(t-1)}}\big\{\mathbf{1}(V_{it\tau}\le s)-\mathbf{1}(V_{it\tau}<0)\big\}ds\Big]$$

$\stackrel{\text{def}}{=} v^\top\xi_1+\xi_2$. According to Kato (2009), in order to prove $\hat{v}$ takes the representation in (4.2), it suffices to prove (a) $\xi_2 \to_p v^\top\Omega_1 v$ with Ω_1 defined in (C3) being a positive definite matrix with uniformly bounded eigenvalues on B, and (b) ξ_1 is tight for $\tau \in B \in (0,1)$, and ξ_1 converges in distribution to a Brownian Bridge. We then prove (a), (b) in the following two steps.

Step a. Proof of (a).

Define $\xi_{2it} = \int_0^{\nu_{i(t-1)}}\big\{\mathbf{1}(V_{it\tau}\le s)-\mathbf{1}(V_{it\tau}<0)\big\}ds$. To prove $\xi_2 = \sum_{i=1}^N\sum_{t=1}^T\xi_{2it} \to_p v^\top\Omega_1 v$, we decompose ξ_{2it} as $\xi_{2it} = \mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1})+\overline{\xi}_{2it}$, where $\overline{\xi}_{2it} = \xi_{2it}-\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1})$. We then prove $\sum_{i=1}^N\sum_{t=1}^T\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1}) \to_p 2^{-1}v^\top\Omega_1 v$ and $\sum_{i=1}^N\sum_{t=1}^T\overline{\xi}_{2it} \to_p 0$ respectively as follows.

We first evaluate $\sum_{i=1}^N\sum_{t=1}^T\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1})$. It can be expressed that $\sum_{i,t}\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1}) = \sum_{i=1}^N\sum_{t=1}^T\mathsf{E}[\int_0^{\nu_{i(t-1)}}\{\mathbf{1}(V_{it\tau}\le s)-\mathbf{1}(V_{it\tau}<0)\}ds|\mathcal{F}_{t-1}] = \sum_{i=1}^N\sum_{t=1}^T\int_0^{\nu_{i(t-1)}}\{F_{i(t-1)}(s+F^{-1}_{i(t-1)}(\tau)) - F_{it-1}(F^{-1}_{i(t-1)}(\tau))\}/s\cdot s ds$. This yields that

$$\begin{aligned}\sum_{i,t}\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1}) &= \sum_{i,t}\int_0^{\nu_{i(t-1)}} f_{it-1}(F^{-1}_{it-1}(\tau))s ds + o_p(1)\\ &= \sum_{i,t}(2NT)^{-1}f_{i(t-1)}\big(X_{i(t-1)}^\top\theta(\tau)\big)v^\top X_{i(t-1)}X_{i(t-1)}^\top v + o_p(1) \to_p 1/2v^\top\Omega_1 v \qquad \text{(B.16)}\end{aligned}$$

according to condition (C3).

Next, we prove $\sum_{i,t}\overline{\xi}_{2it} \to_p 0$. It is not difficult to see that $\overline{\xi}_{2it}$ is a martingale difference sequence, which can be written as $\overline{\xi}_{2it} = \int_0^{\nu_{i(t-1)}}\delta_{it\tau}(s)-\delta_{it\tau}(0)ds$, where

$\delta_{it\tau}(s) = \big\{\mathbf{1}(V_{it\tau} \leq s) - F_{i(t-1)}(s + X_{i(t-1)}^{\top}\theta(\tau))\big\}$. It suffices to show $\mathsf{E}(|\sum_{i,t}\overline{\xi}_{2it}|)^2 = \sum_{i_1,i_2}\sum_{t_1,t_2}\mathsf{E}(\overline{\xi}_{2i_1t_1}\overline{\xi}_{2i_2t_2}) \to 0$. Importantly, recall that $V_{it\tau} = X_{i(t-1)}^{\top}(\theta(U_{it}) - \theta(\tau))$, therefore $V_{it\tau}$ and $V_{jt\tau}$ would be conditionally independent on $\mathcal{F}_{t-1}$. Thus it can be shown that $\mathsf{E}\{\int_0^{\nu_{i(t-1)}}\delta_{it\tau}(s)ds\int_0^{\nu_{j(t-1)}}\delta_{jt\tau}(s)ds\} = \mathsf{E}[\mathsf{E}\{\int_0^{\nu_{i(t-1)}}\delta_{it\tau}(s)ds\int_0^{\nu_{j(t-1)}}\delta_{jt\tau}(s)ds$ $|\mathcal{F}_{t-1}\}] = 0$ due to the conditional independence of $\delta_{it\tau}(s)$ and $\delta_{jt\tau}(s)$ given $\mathcal{F}_{t-1}$. Similarly, for $t_1 > t_2$ we have $\mathsf{E}\{\int_0^{\nu_{i(t_1-1)}}\delta_{it_1\tau}(s)ds\int_0^{\nu_{i(t_2-1)}}\delta_{jt_2\tau}(s)ds\} = \mathsf{E}[\mathsf{E}\{\int_0^{\nu_{i(t_1-1)}}\delta_{it_1\tau}(s)ds$ $\int_0^{\nu_{i(t_2-1)}}\delta_{jt_2\tau}(s)ds|\mathcal{F}_{t_1-1}\}] = 0$. Therefore, we have $\mathsf{E}\{\overline{\xi}_{2i_1t_1\tau}\overline{\xi}_{2i_2t_2\tau}\} = 0$ for $i_1 \neq i_2$ or $t_1 \neq t_2$. Then $\sum_{i_1,i_2}\sum_{t_1,t_2}\mathsf{E}(\overline{\xi}_{2i_1t_1}\overline{\xi}_{2i_2t_2}) = \sum_i\sum_t\mathsf{E}(\overline{\xi}_{2it}^2)$. Next, write $\mathsf{E}(\overline{\xi}_{2it}^2) = \mathsf{E}(\xi_{2it}^2) - \mathsf{E}\{\mathsf{E}(\xi_{2it}|\mathcal{F}_{t-1})\}^2$. Further it can be derived that $\mathsf{E}(\xi_{2it}^2) = \mathsf{E}\,|\int_0^{\nu_{i(t-1)}}\{\mathbf{1}(V_{it\tau} \leq s) - \mathbf{1}(V_{it\tau} \leq 0)\}ds|^2 \leq |\nu_{i(t-1)}|\,\mathsf{E}\int_0^{|\nu_{i(t-1)}|}\{\mathbf{1}(V_{it\tau} \leq s) - \mathbf{1}(V_{it\tau} \leq 0)\}^2 ds$ by the Chebyshev's inequality. Further we have $|\nu_{i(t-1)}|\,\mathsf{E}[\int_0^{|\nu_{i(t-1)}|}\{\mathbf{1}(V_{it\tau} \leq s) - \mathbf{1}(V_{it\tau} \leq 0)\}ds] = |\nu_{i(t-1)}|\,\mathsf{E}[\int_0^{|\nu_{i(t-1)}|}\{F_{i(t-1)}(s + F_{i(t-1)}^{-1}(\tau)) - F_{i(t-1)}(F_{i(t-1)}^{-1}(\tau))\}/s \cdot s ds]$. By similar technique with (B.16), one could obtain $\sum_{i,t}\mathsf{E}(\overline{\xi}_{2it}^2) \leq \mathsf{E}\{\sum_{i,t}2^{-1}(NT)^{-3/2}|f_{i(t-1)}(X_{i(t-1)}^{\top}v)|$ $|X_{i(t-1)}^{\top}v|^{3/2}\} + o(1)$. Since we have $f_{it}(\cdot)$ is bounded and $(NT)^{-3/2}\sum_{i,t}\mathsf{E}(v^{\top}X_{it}X_{it}^{\top}v)^2 = \mathcal{O}\big((NT)^{-1/2}\big) \to 0$, then it can be obtained that $\sum_{i,t}\mathsf{E}(\overline{\xi}_{2it}^2) \to 0$. Lastly, following similar argument of tightness as in Wagener et al. (2012), we can prove that $\sum_{i,t}\xi_{2it} \to_p 0$ uniformly over $\tau \in B$. This completes the proof of $\xi_2 \to_p v^{\top}\Omega_1 v$ for any $\tau \in (0,1)$.

Step b. (Proof of Theorem 4.2)

In this step, we are going to show ξ_1 converges in distribution to a Brownian Bridge $\Omega_0^{1/2}B_{q+3}(\tau)$, where Ω_0 is defined in (C3), and $B_{q+3}(\tau)$ is a $(q+3)$-dimensional Brownian bridge. To prove this conclusion, we adopt two steps:

(B.1) For arbitrary k-dimensional vector $(\tau_1, \tau_2, \cdots, \tau_k)^{\top} \in \mathbb{R}^p$ and $\eta \in \mathbb{R}^{q+3}$, $\big(\xi_1(\tau_1), \xi_1(\tau_2), \cdots, \xi_1(\tau_k)\big)^{\top}\eta \in \mathbb{R}^k$ converge to a k-dimensional multivariate normal distribution.

(B.2) $\eta^{\top}\xi_1(\tau)$ for $\tau \in B \in (0,1)$ is tight, where B is a compact set in $(0,1)$.

Step B.1. Denote $\psi_t = \big(\psi(V_{1t\tau}), \cdots, \psi(V_{Nt\tau})\big)^\top \in \mathbb{R}^N$ for convenience. We then have $\mathsf{E}(\mathbb{X}_{t-1}^\top \psi_t | \mathcal{F}_{t-1}) = 0$. Therefore, $\mathbb{X}_{t-1}^\top \psi_t$ is a martingale difference sequence for $1 \le t \le T$. To prove (B.1), we define $\zeta_t = (NT_N)^{-1/2} \eta^\top \mathbb{X}_{t-1}^\top \psi_t$ and $\mathbb{S}_{Nt} = \sum_{s=1}^t \zeta_{\eta t}$. Then one can see that $\{\zeta_t, \mathcal{F}_{t-1}, -\infty < t < T_N, N \ge 1\}$ is a martingale array, where the number of observed time points T_N is assumed to depend on N with $T_N \to \infty$ as $N \to \infty$. As a result, the double sequence $\{\mathbb{S}_{Nt}, \mathcal{F}_t, -\infty < t \le T_N, N \ge 1\}$ is a martingale array. As a consequence, the martingale difference central limit theorem can be applied (Hall and Heyde, 2014). Specifically, it requires two conditions as follows. First we have

$$\begin{aligned}\sum_{t=1}^{T_N} \mathsf{E}\{\zeta_t^2 \mathbf{1}|\zeta_t^2| > \delta | \mathcal{F}_{t-1}\} &\le \delta^{-2} \sum_{t=1}^{T_N} \mathsf{E}(|\zeta_t|^4 | \mathcal{F}_{t-1}) \\ &\le \delta^{-2} \tau^2 (1-\tau)^2 (NT_N)^{-2} \sum_{t=1}^{T_N} (\eta^\top \mathbb{X}_{t-1}^\top \mathbb{X}_{t-1} \eta)^2 \to_p 0, \end{aligned} \qquad \text{(B.17)}$$

where the last inequality is due to $\mathsf{E}\,\psi^4(V_{it\tau}) \le \tau^2(1-\tau)^2$. Since by the proof of (d) of Lemma 5.1, we have $(NT_N)^{-2} \sum_{t=1}^{T_N} \mathsf{E}(\eta^\top \mathbb{X}_{t-1}^\top \mathbb{X}_{t-1} \eta)^2 \to 0$. Therefore (B.17) can be implied. Secondly, we also have the condition

$$\sum_{t=1}^{T_N} \mathsf{E}\{\zeta_t^2 | \mathcal{F}_{t-1}\} = \frac{\tau(1-\tau)}{NT} \sum_{t=1}^{T_N} \eta^\top \mathbb{X}_{t-1}^\top \mathbb{X}_{t-1} \eta \to_p \tau(1-\tau) \eta^\top \Omega_0 \eta, \qquad \text{(B.18)}$$

by (d) of Lemma 5.1 in Appendix B.1. Therefore, by the central limit theorem for martingale difference sequence in Hall and Heyde (2014), we have that $\xi_1(\tau)$ converge in distribution to Gaussian distribution $\mathsf{N}(0, \tau(1-\tau)\eta^\top \Omega_0 \eta)$ for fixed τ. The conclusion also holds for any finite dimensional vector $(\tau_1, \tau_2, \cdots, \tau_k)^\top$, which proves (B.1).

Step B.2. Then we prove that $\eta^\top \xi_1(\tau)$ for $\tau \in B \in (0,1)$ is tight. The definition of tightness is given as follows.

DEFINITION 5.3. *A process $W_{NT}(\tau)$ is said to be tight if and only if for any $\delta > 0$ there exists a compact set E such that* $\mathbf{sup}_{\tau\in\mathbf{E}}\mathrm{P}(W_{NT}(\tau) \in E) > 1-\delta$.

Define $\psi_1(D) = -(NT)^{-1/2}\sum_{i,t} X_{i(t-1)}\{\psi_{\tau_2}(V_{it\tau_2}) - \psi_{\tau_1}(V_{it\tau_1})\}$ for any interval $D = (\tau_1, \tau_2]$. To show the tightness, we adopt Theorem 15.6 in Billingsley (1968) and prove a sufficient Chentsov-Billingsley type of inequality as follows.

LEMMA 5.2. *For any two intervals $D_1 = (\tau_1, \tau_2]$ and $D_2 = (\tau_2, \tau_3]$, we have*

$$\mathsf{E}\left[\left\{\eta^\top \xi_1(D_1)\right\}^2 \left\{\eta^\top \xi_1(D_2)\right\}^2\right] \le C(\tau_3 - \tau_1), \tag{B.19}$$

where C is a finite positive constant.

To prove Lemma 5.2, we have $\mathsf{E}[\{\eta^\top\xi_1(D_1)\}^2\{\eta^\top\xi_1(D_2)\}^2] = (NT)^{-2}\,\mathsf{E}[\{\sum_{i,t}\eta^\top X_{i(t-1)} \delta_{it}(\tau_1,\tau_2)\}^2\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_2,\tau_3)\}^2]$, where $\delta_{it}(\tau,\tau') = \psi_{\tau'}(V_{it\tau'}) - \psi_\tau(V_{it\tau})$. Next, by Cauchy's inequality, we have $\mathsf{E}[\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_1,\tau_2)\}^2\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_2,\tau_3)\}^2] \le [\mathsf{E}\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_1,\tau_2)\}^4]^{1/2}[\mathsf{E}\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_2,\tau_3)\}^4]^{1/2}$. Since it can be derived $\mathsf{E}\{\delta_{it}(\tau,\tau')|\mathcal{F}_{t-1}\} = 0$, then $\mathsf{E}[\{\eta^\top X_{i_1(t_1-1)}\delta_{i_1t_1}(\tau,\tau')\}\{\eta^\top X_{i_2(t_2-1)}\delta_{i_2t_2}(\tau,\tau')\}\{\eta^\top X_{i_3(t_3-1)} \delta_{i_3t_3}(\tau,\tau')\}\{\eta^\top X_{i_4(t_4-1)}\delta_{i_4t_4}(\tau,\tau')\}]$ is non-zero only if (a) $i_1 = i_2, t_1 = t_2$ and $i_3 = i_4 \neq i_1, t_3 = t_4 \neq t_1$ or (b) $i_1 = i_2 = i_3 = i_4$ and $t_1 = t_2 = t_3 = t_4$. It is straightforward to verify $(NT)^{-2}\,\mathsf{E}\{\sum_{i,t}\eta^\top X_{i(t-1)}\delta_{it}(\tau_1,\tau_2)\}^4 =$

$$(NT)^{-2}\Big[\sum_{i,t}\mathsf{E}\left\{(\eta^\top X_{i(t-1)})^2\delta_{it}^2(\tau_1,\tau_2)\right\}\Big]^2 + (NT)^{-2}\sum_{i,t}\mathsf{E}\left\{(\eta^\top X_{i(t-1)})^4\delta_{it}^4(\tau_1,\tau_2)\right\}.$$

By the proof of (d) in Lemma 5.1, we know that $\mathsf{E}(\eta^\top X_{it})^2 = \mathcal{O}(1)$ and $\mathsf{E}(\eta^\top X_{it})^4 = \mathcal{O}(1)$. Moreover, it can be verified $\mathsf{E}\{\delta_{it}^2(\tau_1,\tau_2)\} \le \tau_2 - \tau_1$ and $\mathsf{E}\{\delta_{it}^4(\tau_1,\tau_2)\} \le \tau_2 - \tau_1$.

By combining the results together, we have

$$\mathsf{E}\left[\left\{\eta^\top\xi_1(D_1)\right\}^2\left\{\eta^\top\xi_1(D_2)\right\}^2\right] \leq C(\tau_2-\tau_1)(\tau_3-\tau_2) \leq C|\tau_3-\tau_1|,$$

for some positive constant C. This completes the proof of Lemma 5.2. We then conclude that the $\xi_1(\tau)$ converge weakly to a $(q+3)$-dimensional Brownian bridge. Consequently, the Theorem 4.2 can be proved.

References

Koenker, R.; Hallock, K. F. Journal of Econometric Perspectives **2001**, 15, 143–156.

Fitzenberger, B.; Koenker, R.; Machado, J. A. Economic applications of quantile regression; Springer Science & Business Media, 2013.

Gaglianone, W. P.; Lima, L. R.; Linton, O.; Smith, D. R. Journal of Business & Economic Statistics **2012**,

Härdle, W. K.; Wang, W.; Yu, L. Journal of Econometrics **2016**,

Sankarasubramanian, A.; Lall, U. Water Resources Research **2003**, 39.

Wang, W.; Bobojonov, I.; Härdle, W.; Odening, M. Stochastic environmental research and risk assessment **2013**, 27, 1565–1574.

Koenker, R.; Bassett, G. W. Econometrica **1978**, 46, 33–50.

Portnoy, S. Journal of Multivariate Analysis **1991**, 38, 100–113.

Portnoy, S. The Annals of Statistics **1997**, 414–434.

Koenker, R. Quantile regression; Cambridge university press, 2005.

Hallin, M.; Jureveccková, J.; Picek, J.; Zahaf, T. Journal of statistical planning and inference **1999**, 75, 319–330.

Hasan, M. N.; Koenker, R. W. Econometrica: Journal of the Econometric Society **1997**, 133–161.

Engle, R. F.; Manganelli, S. Journal of Business and Economic Statistics **2004**, 22, 367–381.

Koenker, R.; Xiao, Z. Journal of the American Statistical Association **2006**, 101, 980–990.

White, H.; Kim, T.-H.; Manganelli, S. **2010**,

Baruník, J.; Kley, T. Available at SSRN 2678977 **2015**,

Sewell, D. K.; Chen, Y. Journal of the American Statistical Association **2015**, 110, 1646–1657.

Zhao, Y.; Levina, E.; Zhu, J. Proceedings of the National Academy of Sciences **2011**, 108, 7321–7326.

others,, et al. The Annals of Statistics **2013**, 41, 2097–2122.

others,, et al. Bayesian Analysis **2016**,

Bickel, P. J.; Chen, A. Proceedings of the National Academy of Sciences **2009**, 106, 21068–21073.

Zhao, Y.; Levina, E.; Zhu, J. The Annals of Statistics **2012**, 2266–2292.

Zhang, J.; Chen, Y. Journal of the American Statistical Association **2013**, 108, 1295–1307.

Zhou, J.; Tu, Y.; Chen, Y.; Wang, H. Journal of Business & Economic Statistics **2015**, To appear.

Huang, L. W. Z. H., D.; Wang, H. Working Paper **2016**,

Zhu, X.; Pan, R.; Li, G.; Liu, Y.; Wang, H. Annals of statistics **2016**,

Carrington, P. J.; Scott, J.; Wasserman, S. Models and methods in social network analysis; Cambridge university press, 2005; Vol. 28.

Newman, M. Networks: an introduction; Oxford university press, 2010.

Kolaczyk, E. D.; Csárdi, G. Statistical analysis of network data with R; Springer, 2014.

Abello, J.; Pardalos, P. M.; Resende, M. G. Handbook of massive data sets; Springer, 2013; Vol. 4.

Li, G.; Li, Y.; Tsai, C.-L. Journal of the American Statistical Association **2015**, 110, 246–261.

Lütkepohl, H. New introduction to multiple time series analysis; Springer Science & Business Media, 2005.

Chen, X.; Chen, Y.; Xiao, P. Journal of Marketing Research **2013**, 50, 95–110.

Fan, J.; Fan, Y. Manuscript **2010**,

Holland, P. W.; Leinhardt, S. Journal of the american Statistical association **1981**, 76, 33–50.

Wang, Y. J.; Wong, G. Y. Journal of the American Statistical Association **1987**, 82, 8–19.

Nowicki, K.; Snijders, T. A. B. Journal of the American Statistical Association **2001**, 96, 1077–1087.

Barabási, A.-L.; Albert, R. science **1999**, 286, 509–512.

Clauset, A.; Shalizi, C. R.; Newman, M. E. SIAM review **2009**, 51, 661–703.

Fama, E. F.; French, K. R. Journal of Financial Economics **2015**, 116, 1–22.

Banerjee, S.; Carlin, B. P.; Gelfand, A. E. Hierarchical modeling and analysis for spatial data; Crc Press, 2014.

Bardet, J.-M.; Doukhan, P.; Lang, G.; Ragache, N. ESAIM: Probability and Statistics **2008**, 12, 154–172.

Andrews, D. W. Econometric theory **1988**, 4, 458–467.

Gallant, A. R. Nonlinear statistical models; John Wiley & Sons, 2009; Vol. 310.

Kato, K. Journal of Multivariate Analysis **2009**, 100, 1816–1829.

Wagener, J.; Volgushev, S.; Dette, H. Mathematical Methods of Statistics **2012**, 21, 127–141.

Hall, P.; Heyde, C. C. Martingale limit theory and its application; Academic press, 2014.

Billingsley, P. Convergence of Probability Measures; Wiley, New York, 1968.

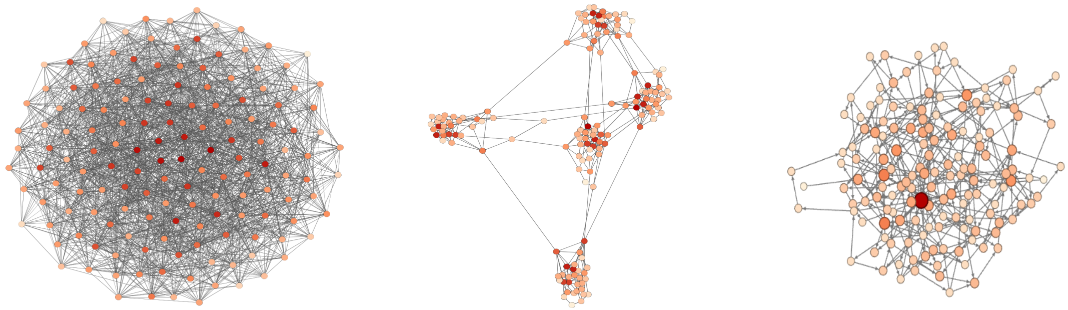

Figure 1: The left panel: dyad independence network; The middle panel: stochastic block model; the right panel: power-law distribution network.

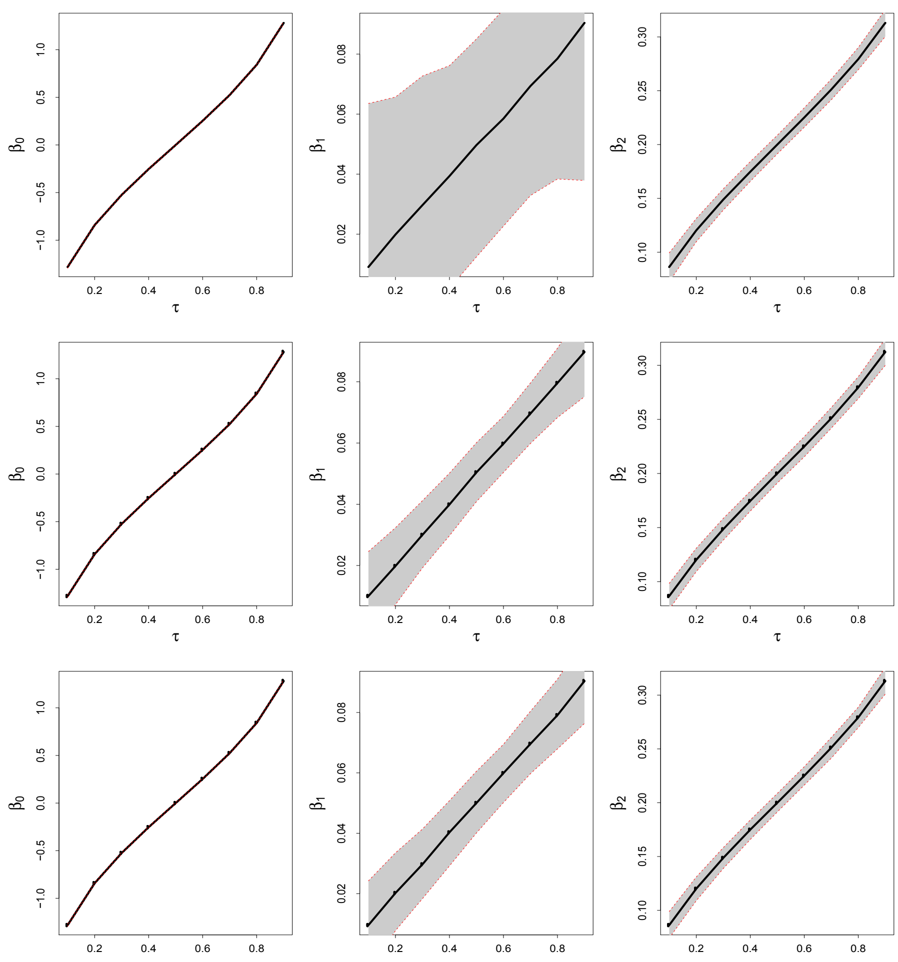

Figure 2: The estimated β_0 to β_2 against τ. The top panel: dyad independence network; The middle panel: stochastic block model; the bottom panel: power-law distribution network.

Table 1: Simulation Results for dyad independence network with 1000 Replications. The RMSE ($\times 10^{-2}$) and the Coverage Probability (%) are reported for β_0 to β_1. The RMSE is also reported for γ. Lastly, the network density is computed and given. Z stands for normal distribution and T stands for t- distribution

N	Dist.	β_0	β_1	β_2	γ	ND
			$\tau = 0.1$			
100	Z	2.60(95.0)	10.10(95.8)	2.47(94.3)	3.09	22.7
	T	3.43(96.4)	11.22(95.2)	2.37(95.6)	4.17	
500	Z	1.08(96.2)	4.61(95.4)	1.04(96.0)	1.32	4.7
	T	1.51(95.4)	4.90(95.9)	1.03(96.1)	1.82	
1000	Z	0.77(95.8)	3.29(95.0)	0.80(94.0)	0.93	2.4
	T	1.06(95.8)	3.66(95.0)	0.75(95.0)	1.29	
			$\tau = 0.5$			
100	Z	1.90(95.5)	6.62(95.4)	1.65(96.7)	2.11	22.7
	T	1.99(95.7)	5.67(94.5)	1.32(93.3)	2.15	
500	Z	0.84(94.4)	2.99(95.5)	0.79(94.9)	0.87	4.7
	T	0.90(94.9)	2.43(96.2)	0.55(92.3)	0.91	
1000	Z	0.59(94.7)	2.17(95.0)	0.53(95.7)	0.63	2.4
	T	0.62(94.2)	1.77(95.0)	0.37(93.5)	0.66	
			$\tau = 0.9$			
100	Z	2.57(95.3)	9.96(95.1)	2.49(94.1)	2.92	22.7
	T	3.61(95.0)	10.61(95.4)	2.41(94.5)	3.98	
500	Z	1.08(96.3)	4.27(95.8)	1.10(94.0)	1.30	4.7
	T	1.53(95.6)	4.75(94.8)	1.11(93.9)	1.75	
1000	Z	0.78(95.5)	3.14(95.5)	0.76(95.0)	0.90	2.4
	T	1.09(95.9)	3.41(96.0)	0.84(93.5)	1.26	

Table 2: Simulation Results for stochastic block network with 1000 Replications. The RMSE ($\times 10^{-2}$) and the Coverage Probability (%) are reported for β_0 to β_1. The RMSE is also reported for γ. Lastly, the network density is computed and given.

N	Dist.	β_0	β_1	β_2	γ	ND
			$\tau = 0.1$			
100	Z	2.61(95.8)	3.29(94.9)	2.45(94.3)	3.03	2.6
	T	3.33(96.7)	3.37(96.0)	2.40(94.2)	4.29	
500	Z	1.14(94.3)	1.40(94.5)	1.08(94.9)	1.32	0.5
	T	1.57(94.0)	1.50(95.1)	1.04(95.6)	1.82	
1000	Z	0.79(94.6)	0.89(95.0)	0.74(95.9)	0.94	0.2
	T	1.09(95.4)	0.95(94.9)	0.78(94.5)	1.28	
			$\tau = 0.5$			
100	Z	1.88(94.5)	2.15(94.2)	1.74(95.2)	2.07	2.6
	T	2.03(94.0)	1.76(95.1)	1.28(93.4)	2.17	
500	Z	0.84(94.5)	0.92(94.5)	0.77(94.9)	0.90	0.5
	T	0.86(94.7)	0.75(94.5)	0.52(93.2)	0.90	
1000	Z	0.59(94.4)	0.59(95.9)	0.53(95.6)	0.63	0.2
	T	0.61(95.4)	0.47(95.6)	0.38(93.0)	0.64	
			$\tau = 0.9$			
100	Z	2.56(95.0)	2.91(96.0)	2.46(94.5)	2.94	2.6
	T	3.44(95.8)	3.28(94.3)	2.39(94.3)	4.07	
500	Z	1.08(95.4)	1.33(94.6)	1.07(95.3)	1.29	0.5
	T	1.52(95.9)	1.45(95.8)	1.12(94.0)	1.78	
1000	Z	0.80(95.2)	0.89(94.4)	0.75(96.0)	0.91	0.2
	T	1.03(96.4)	0.90(95.3)	0.82(93.4)	1.23	

Table 3: Simulation Results for power-law distribution network with 1000 Replications. The RMSE ($\times 10^{-2}$) and the Coverage Probability (%) are reported for β_0 to β_1. The RMSE is also reported for γ. Lastly, the network density is computed and given.

N	Dist.	β_0	β_1	β_2	γ	ND
			$\tau = 0.1$			
100	Z	2.44(95.9)	2.95(95.4)	2.32(96.2)	3.08	2.4
	T	3.45(96.3)	3.28(93.9)	2.36(95.1)	4.19	
500	Z	1.09(95.5)	1.24(96.3)	1.07(95.4)	1.35	0.5
	T	1.53(94.7)	1.42(94.8)	1.04(96.2)	1.79	
1000	Z	0.76(95.8)	0.91(95.6)	0.77(94.7)	0.94	0.2
	T	1.06(95.0)	0.99(95.5)	0.75(95.3)	1.28	
			$\tau = 0.5$			
100	Z	1.87(95.3)	1.96(95.7)	1.79(94.4)	2.07	2.4
	T	1.94(96.4)	1.55(96.2)	1.29(93.1)	2.15	
500	Z	0.82(95.7)	0.85(94.6)	0.77(95.8)	0.89	0.5
	T	0.90(95.5)	0.71(94.0)	0.54(93.3)	0.92	
1000	Z	0.58(95.1)	0.62(94.2)	0.54(96.0)	0.62	0.2
	T	0.63(94.4)	0.51(92.2)	0.37(94.6)	0.64	
			$\tau = 0.9$			
100	Z	2.55(95.8)	2.94(93.5)	2.43(94.3)	2.91	2.4
	T	3.53(95.3)	3.01(94.7)	2.43(94.1)	4.11	
500	Z	1.12(95.1)	1.20(96.3)	1.09(95.1)	1.29	0.5
	T	1.51(95.5)	1.33(95.1)	1.10(94.3)	1.80	
1000	Z	0.79(95.2)	0.87(95.7)	0.76(95.1)	0.90	0.2
	T	1.09(94.7)	0.98(95.3)	0.83(92.1)	1.26	

Table 4: The detailed NQAR analysis results for the Stock dataset (τ =0.05, 0.5, 0.95). The yearly estimates ($\times 10^{-2}$) are reported with the standard error ($\times 10^{-2}$) given in parentheses. The p-values are also reported.

	$\tau = 0.05$		$\tau = 0.5$		$\tau = 0.95$	
	Est.	p-value	Est.	p-value	Est.	p-value
$\hat{\beta}_0$	0.05 (0.00)	< 0.01	1.00 (0.04)	< 0.01	2.96 (0.13)	< 0.01
$\hat{\beta}_1$	0.00 (0.02)	0.99	-0.04 (0.77)	0.95	6.09 (2.16)	< 0.01
$\hat{\beta}_2$	4.16 (0.14)	< 0.01	35.70 (0.47)	< 0.01	67.84 (1.13)	< 0.01
SIZE	0.00 (0.01)	0.98	-1.00 (0.09)	< 0.01	-4.10 (0.28)	< 0.01
BM	0.00 (0.01)	0.99	-0.29 (0.04)	< 0.01	-0.71 (0.25)	< 0.01
PR	0.00 (0.00)	1.00	-0.30 (0.12)	0.01	0.39 (0.38)	0.31
AR	-0.02 (0.03)	0.53	-0.66 (0.11)	< 0.01	-0.47 (0.36)	0.20
CASH	-0.01 (0.01)	0.03	-0.14 (0.06)	0.01	-0.05 (0.27)	0.86
LEV	0.00 (0.01)	0.97	-0.79 (0.05)	< 0.01	-2.42 (0.44)	< 0.01

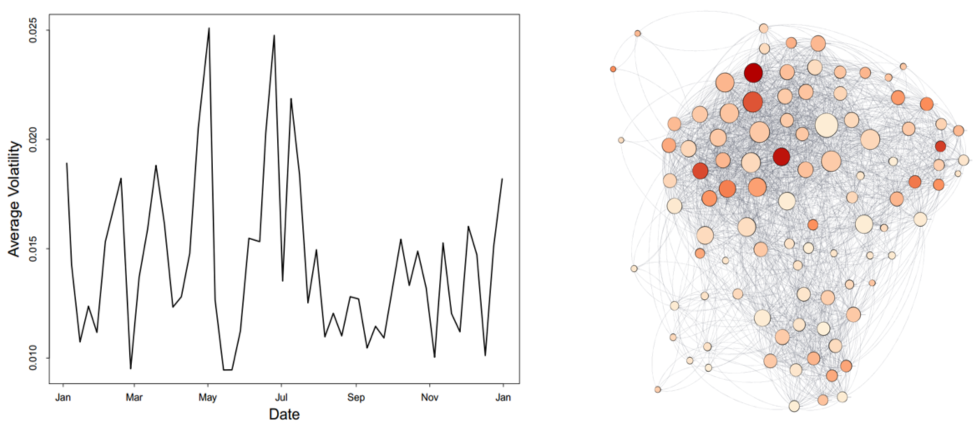

Figure 3: The left panel: the average stock volatility of Chinese A stock market in 2013; the right panel: the common shareholder network of top 100 market value stocks in 2013.

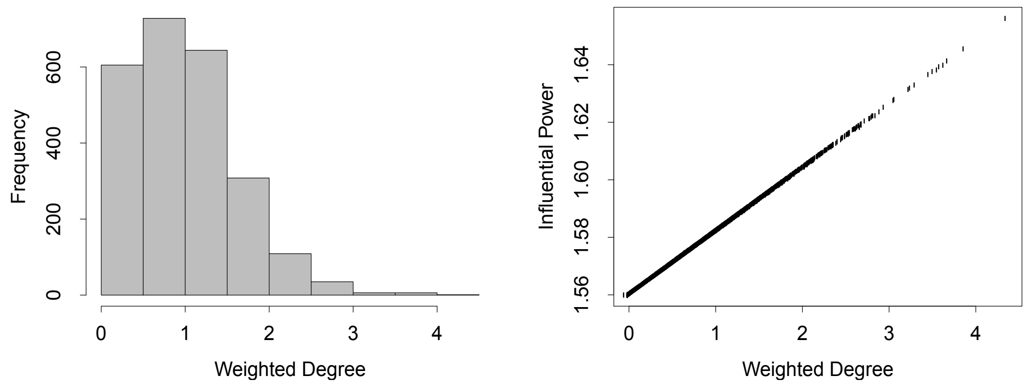

Figure 4: The left panel: the histogram of the weighted degrees; the right panel: the influential power against weighted degrees.

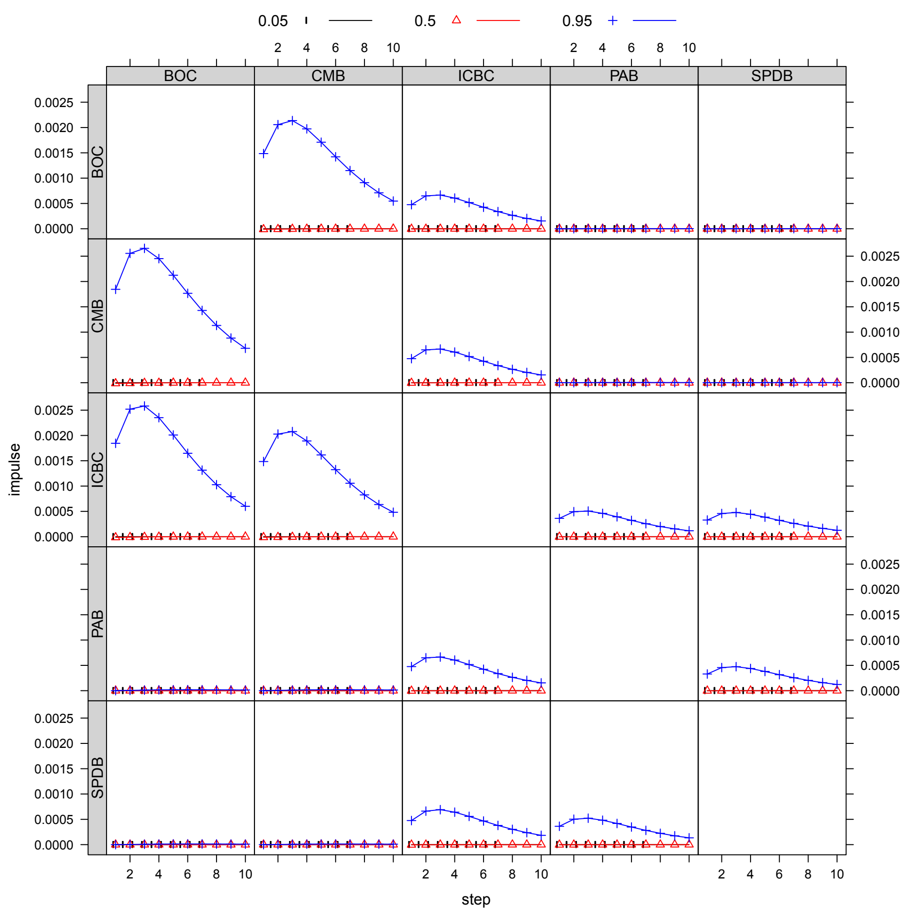

Figure 5: Impulse analysis for $\tau = 0.05, 0.5, 0.95$. The cross-sectional impulse effect intensity between BOC, CMB, ICBC, PAB, and SPDB are given. The impulse direction is from column to row.

SFB 649 Discussion Paper Series 2016

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Downside risk and stock returns: An empirical analysis of the long-run and short-run dynamics from the G-7 Countries" by Cathy Yi-Hsuan Chen, Thomas C. Chiang and Wolfgang Karl Härdle, January 2016.
002 "Uncertainty and Employment Dynamics in the Euro Area and the US" by Aleksei Netsunajev and Katharina Glass, January 2016.
003 "College Admissions with Entrance Exams: Centralized versus Decentralized" by Isa E. Hafalir, Rustamdjan Hakimov, Dorothea Kübler and Morimitsu Kurino, January 2016.
004 "Leveraged ETF options implied volatility paradox: a statistical study" by Wolfgang Karl Härdle, Sergey Nasekin and Zhiwu Hong, February 2016.
005 "The German Labor Market Miracle, 2003 -2015: An Assessment" by Michael C. Burda, February 2016.
006 "What Derives the Bond Portfolio Value-at-Risk: Information Roles of Macroeconomic and Financial Stress Factors" by Anthony H. Tu and Cathy Yi-Hsuan Chen, February 2016.
007 "Budget-neutral fiscal rules targeting inflation differentials" by Maren Brede, February 2016.
008 "Measuring the benefit from reducing income inequality in terms of GDP" by Simon Voigts, February 2016.
009 "Solving DSGE Portfolio Choice Models with Asymmetric Countries" by Grzegorz R. Dlugoszek, February 2016.
010 "No Role for the Hartz Reforms? Demand and Supply Factors in the German Labor Market, 1993-2014" by Michael C. Burda and Stefanie Seele, February 2016.
011 "Cognitive Load Increases Risk Aversion" by Holger Gerhardt, Guido P. Biele, Hauke R. Heekeren, and Harald Uhlig, March 2016.
012 "Neighborhood Effects in Wind Farm Performance: An Econometric Approach" by Matthias Ritter, Simone Pieralli and Martin Odening, March 2016.
013 "The importance of time-varying parameters in new Keynesian models with zero lower bound" by Julien Albertini and Hong Lan, March 2016.
014 "Aggregate Employment, Job Polarization and Inequalities: A Transatlantic Perspective" by Julien Albertini and Jean Olivier Hairault, March 2016.
015 "The Anchoring of Inflation Expectations in the Short and in the Long Run" by Dieter Nautz, Aleksei Netsunajev and Till Strohsal, March 2016.
016 "Irrational Exuberance and Herding in Financial Markets" by Christopher Boortz, March 2016.
017 "Calculating Joint Confidence Bands for Impulse Response Functions using Highest Density Regions" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, March 2016.
018 "Factorisable Sparse Tail Event Curves with Expectiles" by Wolfgang K. Härdle, Chen Huang and Shih-Kang Chao, March 2016.
019 "International dynamics of inflation expectations" by Aleksei Netšunajev and Lars Winkelmann, May 2016.
020 "Academic Ranking Scales in Economics: Prediction and Imdputation" by Alona Zharova, Andrija Mihoci and Wolfgang Karl Härdle, May 2016.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2016

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

021 "CRIX or evaluating blockchain based currencies" by Simon Trimborn and Wolfgang Karl Härdle, May 2016.
022 "Towards a national indicator for urban green space provision and environmental inequalities in Germany: Method and findings" by Henry Wüstemann, Dennis Kalisch, June 2016.
023 "A Mortality Model for Multi-populations: A Semi-Parametric Approach" by Lei Fang, Wolfgang K. Härdle and Juhyun Park, June 2016.
024 "Simultaneous Inference for the Partially Linear Model with a Multivariate Unknown Function when the Covariates are Measured with Errors" by Kun Ho Kim, Shih-Kang Chao and Wolfgang K. Härdle, August 2016.
025 "Forecasting Limit Order Book Liquidity Supply-Demand Curves with Functional AutoRegressive Dynamics" by Ying Chen, Wee Song Chua and Wolfgang K. Härdle, August 2016.
026 "VAT multipliers and pass-through dynamics" by Simon Voigts, August 2016.
027 "Can a Bonus Overcome Moral Hazard? An Experiment on Voluntary Payments, Competition, and Reputation in Markets for Expert Services" by Vera Angelova and Tobias Regner, August 2016.
028 "Relative Performance of Liability Rules: Experimental Evidence" by Vera Angelova, Giuseppe Attanasi, Yolande Hiriart, August 2016.
029 "What renders financial advisors less treacherous? On commissions and reciprocity" by Vera Angelova, August 2016.
030 "Do voluntary payments to advisors improve the quality of financial advice? An experimental sender-receiver game" by Vera Angelova and Tobias Regner, August 2016.
031 "A first econometric analysis of the CRIX family" by Shi Chen, Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle, TM Lee and Bobby Ong, August 2016.
032 "Specification Testing in Nonparametric Instrumental Quantile Regression" by Christoph Breunig, August 2016.
033 "Functional Principal Component Analysis for Derivatives of Multivariate Curves" by Maria Grith, Wolfgang K. Härdle, Alois Kneip and Heiko Wagner, August 2016.
034 "Blooming Landscapes in the West? - German reunification and the price of land." by Raphael Schoettler and Nikolaus Wolf, September 2016.
035 "Time-Adaptive Probabilistic Forecasts of Electricity Spot Prices with Application to Risk Management." by Brenda López Cabrera , Franziska Schulz, September 2016.
036 "Protecting Unsophisticated Applicants in School Choice through Information Disclosure" by Christian Basteck and Marco Mantovani, September 2016.
037 "Cognitive Ability and Games of School Choice" by Christian Basteck and Marco Mantovani, Oktober 2016.
038 "The Cross-Section of Crypto-Currencies as Financial Assets: An Overview" by Hermann Elendner, Simon Trimborn, Bobby Ong and Teik Ming Lee, Oktober 2016.
039 "Disinflation and the Phillips Curve: Israel 1986-2015" by Rafi Melnick and Till Strohsal, Oktober 2016.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2016

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

040 "Principal Component Analysis in an Asymmetric Norm" by Ngoc M. Tran, Petra Burdejová, Maria Osipenko and Wolfgang K. Härdle, October 2016.
041 "Forward Guidance under Disagreement - Evidence from the Fed's Dot Projections" by Gunda-Alexandra Detmers, October 2016.
042 "The Impact of a Negative Labor Demand Shock on Fertility - Evidence from the Fall of the Berlin Wall" by Hannah Liepmann, October 2016.
043 "Implications of Shadow Bank Regulation for Monetary Policy at the Zero Lower Bound" by Falk Mazelis, October 2016.
044 "Dynamic Contracting with Long-Term Consequences: Optimal CEO Compensation and Turnover" by Suvi Vasama, October 2016.
045 "Information Acquisition and Liquidity Dry-Ups" by Philipp Koenig and David Pothier, October 2016.
046 "Credit Rating Score Analysis" by Wolfgang Karl Härdle, Phoon Kok Fai and David Lee Kuo Chuen, November 2016.
047 "Time Varying Quantile Lasso" by Lenka Zbonakova, Wolfgang Karl Härdle, Phoon Kok Fai and Weining Wang, November 2016.
048 "Unraveling of Cooperation in Dynamic Collaboration" by Suvi Vasama, November 2016.
049 "Q3-D3-LSA" by Lukas Borke and Wolfgang K. Härdle, November 2016.
050 "Network Quantile Autoregression" by Xuening Zhu, Weining Wang, Hangsheng Wang and Wolfgang Karl Härdle, November 2016.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".